

INFORMATION FOR SEQ ID NO: 173:

SEQUENCE CHARACTERISTICS:

LENGTH: 4029 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY:

LOCATION: 1..4029

OTHER INFORMATION: /note= "preliminary sequence for

human TRT cDNA insert of

plasmid pGRN121"

OTHER INFORMATION: plasmid pGRN121"

US-08-854-050-173

Query Match 1.6%; Score 22; DB 4; Length 4029;

Best Local Similarity 100.0%; Pred. No. 1.6;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1360 TTTTGAAGAAAAA 1381

Db 3999 TTTTGAAGAAAAA 4020

RESULT 26

US-09-430-323-173

Sequence 173, Application US/09430323

Patent No. 6309867

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin

Andrews, William H.

TITLE OF INVENTION: No. 6309867el Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/430,323

FILING DATE: 29-Oct-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-0029300S

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 173:

SEQUENCE CHARACTERISTICS:

LENGTH: 4029 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY:

LOCATION: 1..4029

OTHER INFORMATION: /note= "preliminary sequence for

human TRT cDNA insert of

plasmid pGRN121"

OTHER INFORMATION: SEQ ID NO: 173;

US-09-430-323-173

Query Match 1.6%; Score 22; DB 4; Length 4029;

Best Local Similarity 100.0%; Pred. No. 1.6;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1360 TTTTGAAGAAAAA 1381

Db 3999 TTTTGAAGAAAAA 4020

RESULT 27

US-09-306-290-41/c

Sequence 41, Application US/09306290

Patent No. 6221535

GENERAL INFORMATION:

APPLICANT: Rovera, Giovanni

APPLICANT: Mukhopadhyay, Sunil

TITLE OF INVENTION: METHODS FOR SOLID-PHASE AMPLIFICATION OF DNA TEMPLATE

FILE OF INVENTION: (SPADT) USING MULTIRAYS

FILE REFERENCE: 09924-10

CURRENT APPLICATION NUMBER: US/09/306,290

CURRENT FILING DATE: 1999-05-06

NUMBER OF SEQ ID NOS: 43

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 41

LENGTH: 40

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Primer

OTHER INFORMATION: HHV8RF1191

US-09-306-290-41

Query Match 1.5%; Score 21; DB 4; Length 40;

Best Local Similarity 100.0%; Pred. No. 6.1;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1361 TTTGAAAAA 1381

Db 25 TTTGAAAAA 5

RESULT 28

US-09-227-357-121

Sequence 121, Application US/09227357

Patent No. 6342581

GENERAL INFORMATION:

APPLICANT: Fischer et al.

TITLE OF INVENTION: 123 Human Secreted Proteins

FILE REFERENCE: P2010P1

CURRENT APPLICATION NUMBER: US/09/227,357

CURRENT FILING DATE: 1999-01-08

EARLIER APPLICATION NUMBER: PCT/US98/13684

EARLIER FILING DATE: 1998-07-07

EARLIER APPLICATION NUMBER: 60/051,926

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/052,793

EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,925


```
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 121
; LENGTH: 360
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (340)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-227-357-121
```

```
Query Match 1.5%; Score 21; DB 4; Length 360;
Best Local Similarity 100.0%; Pred. No. 5.1;
```

```
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1361 TTGTGAAAAAAAAAAAAAAAA 1381
| | | | | | | | | | | | | | | |
DB 306 TTGTGAAAAAAAAAAAAAAAA 326

RESULT 29
US-08-485-284A-2/c
; Sequence 2, Application US/08485284A
; Patent No. 5750372
; GENERAL INFORMATION:
; APPLICANT: SAKAI, YASUYOSHI
; APPLICANT: TANI, YOSHIKI
; APPLICANT: SHIBANO, YUJI
; APPLICANT: KONDO, HIROTO
; APPLICANT: HATANAKA, HARUYO
; TITLE OF INVENTION: NOVEL VECTOR HAVING PROMOTER THAT IS
; INDUCIBLE BY METHANOL AND/OR GLYCEROL
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20003-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,284A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 43361/1992
; FILING DATE: 28-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/025,416
; FILING DATE: 01-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WHITE JR., PAUL E.
; REGISTRATION NUMBER: 32,011
; REFERENCE/DOCKET NUMBER: 217755/FPS382090S
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 518 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: genomic DNA
US-08-485-284A-2

Query Match 1.5%; Score 21; DB 1; Length 518;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1361 TTGTGAAAAAAAAAAAAAAAA 1381
| | | | | | | | | | | | | | | |
DB 165 TTGTGAAAAAAAAAAAAAAAA 145

RESULT 30
US-07-846-992-1
; Sequence 1, Application US/07846992
; Patent No. 5583046
; GENERAL INFORMATION:
```

```
/ APPLICANT: Valenta, Rudolf
/ APPLICANT: Duchene, Michael
/ APPLICANT: Petteburger, Karin
/ APPLICANT: Breitenbach, Michael
/ APPLICANT: Kraft, Dietrich
/ APPLICANT: Rumpold, Helmut
/ APPLICANT: Scheiner, Otto
/ TITLE OF INVENTION: Birch Pollen Allergen p14 for Diagnosis
/ TITLE OF INVENTION: and Therapy of Allergic Diseases
/ NUMBER OF SEQUENCES: 11
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Pennie & Edmonds
/ STREET: 1155 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036-2711
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/846,992
/ FILING DATE: 19920606
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/353,844
/ FILING DATE: 18-MAY-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Jones III, Harry C
/ REGISTRATION NUMBER: 20,280
/ REFERENCE/DOCKET NUMBER: 6530-011
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-8864/9741
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 700 base pairs
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Betula verrucosa
/ IMMEDIATE SOURCE:
/ US-07-846-992-1

Query Match 1.5%; Score 21; DB 1; Length 700;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTGTGAAAAAIAAAAAAAAAA 1381
DB 657 TTGTGAAAAAIAAAAAAAAAA 677

RESULT 31
US-08-469-555-1
/ Sequence 1, Application US/08469555
/ Patent No. 5648242
/ GENERAL INFORMATION:
/ APPLICANT: Valenta, Rudolf
/ APPLICANT: Duchene, Michael
/ APPLICANT: Petteburger, Karin
/ APPLICANT: Breitenbach, Michael
/ APPLICANT: Kraft, Dietrich
/ APPLICANT: Rumpold, Helmut
/ APPLICANT: Scheiner, Otto
/ TITLE OF INVENTION: Birch Pollen Allergen p14 for Diagnosis
```

```
/ TITLE OF INVENTION: and Therapy of Allergic Diseases
/ NUMBER OF SEQUENCES: 11
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Pennie & Edmonds
/ STREET: 1155 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036-2711
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/469,555
/ FILING DATE: 06-JUN-1995
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/846,992
/ FILING DATE: 06-JUN-1992
/ APPLICATION NUMBER: US/07/353,844
/ FILING DATE: 18-MAY-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Jones III, Harry C
/ REGISTRATION NUMBER: 20,280
/ REFERENCE/DOCKET NUMBER: 6530-011
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-8864/9741
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 700 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Betula verrucosa
/ IMMEDIATE SOURCE:
/ LIBRARY: POLLEN FROM ALLERGEN AB, ENGELHOLM, SWEDEN
/ US-08-469-555-1

Query Match 1.5%; Score 21; DB 1; Length 700;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTGTGAAAAAIAAAAAAAAAA 1381
DB 657 TTGTGAAAAAIAAAAAAAAAA 677

RESULT 32
US-08-946-026-14
/ Sequence 14, Application US/08946026
/ Patent No. 6034218
/ GENERAL INFORMATION:
/ APPLICANT: Reed, Steven G.
/ APPLICANT: Dillon, Davin C.
/ APPLICANT: Twardzik, Daniel R.
/ APPLICANT: Mitcham, Jennifer L.
/ TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
/ TITLE OF INVENTION: AND IMMUNODIAGNOSIS OF PROSTATE CANCER
/ NUMBER OF SEQUENCES: 59
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: SEED and BERRY LLP
/ STREET: 6300 Columbia Center, 701 Fifth Avenue
/ CITY: Seattle
/ STATE: Washington
/ COUNTRY: USA
```

ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/946,026
FILING DATE: 07-OCT-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.424C1
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-946-026-14

Query Match 1.5%: Score 21; DB 3; Length 720;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTTGAAAAA 1381
Db 700 TTTGAAAAA 720

RESULT 33
5185441-40
PATENT NO. 5185441
APPLICANT: WALLNER, BARBARA P.; HESSE, CATHERINE
TITLE OF INVENTION: DNA SEQUENCES, RECOMBINANT DNA
MOLECULES AND PROCESSES FOR PRODUCING PI-LINKED LYMPHOCYTE
FUNCTION ASSOCIATED ANTIGEN-3
NUMBER OF SEQUENCES: 41
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/237,309
FILING DATE: 26-AUG-1986
SEQ ID NO: 40
LENGTH: 855
5185441-40

Query Match 1.5%: Score 21; DB 6; Length 855;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTTGAAAAA 1381
Db 817 TTTGAAAAA 837

RESULT 34
5223394-3
PATENT NO. 5223394
APPLICANT: WALLNER, BARBARA
TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
LINKAGE SIGNAL SEQUENCE
NUMBER OF SEQUENCES: 12
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/335,688
FILING DATE: 10-APR-1989
SEQ ID NO: 3
LENGTH: 855
5223394-3

Query Match 1.5%: Score 21; DB 6; Length 855;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTTGAAAAA 1381
Db 817 TTTGAAAAA 837

RESULT 35
US-07-940-861-11
SEQUENCE 11, Application US/07940861
PATENT NO. 3547853
GENERAL INFORMATION:
APPLICANT: BIOGEN, INC.
APPLICANT: WALLNER, Barbara P.
APPLICANT: MILLER, Glenn T.
APPLICANT: ROSA, Margaret D.
TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 875 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022-6250
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/940,861
FILING DATE: 21-OCT-1992
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02050
FILING DATE: 12-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/667,971
FILING DATE: 12-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,967
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F., Jr.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B151CIP2
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 863 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 18..737
US-07-940-861-11

Query Match 1.5%: Score 21; DB 1; Length 863;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTTGAAAAA 1381
Db 817 TTTGAAAAA 837

Db 825 TTTGAAAAAAAAAAAAAAAAAAAA 845

RESULT 36

US-08-459-512-11
; Sequence 11, Application US/08459512
; Patent No. 5728677

GENERAL INFORMATION:

APPLICANT: BIOGEN, INC.
APPLICANT: WALLNER, Barbara P.
APPLICANT: MILLER, Glenn T.
APPLICANT: ROSA, Margaret D.
TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 875 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022-6250
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/459,512
FILING DATE: 02-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02050
FILING DATE: 12-MAR-1992
APPLICATION NUMBER: US 07/667,971
FILING DATE: 12-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,967
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F., Jr.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B151CIP2
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 863 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 18..737
US-08-459-512-11

Query Match 1.5%; Score 21; DB 1; Length 863;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Qy 1361 TTTGAAAAAAAAAAAAAAAAAAAA 1381

|||||

Db 825 TTTGAAAAAAAAAAAAAAAAAAAA 845

RESULT 37

US-08-459-657-11
; Sequence 11, Application US/08459657
; Patent No. 5914111

GENERAL INFORMATION:
APPLICANT: BIOGEN, INC.
APPLICANT: WALLNER, Barbara P.
APPLICANT: MILLER, Glenn T.
APPLICANT: ROSA, Margaret D.
TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 875 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022-6250
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/459,657
FILING DATE: 02-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02050
FILING DATE: 12-MAR-1992
APPLICATION NUMBER: US 07/667,971
FILING DATE: 12-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,967
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F., Jr.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B151CIP2
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 863 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 18..737
US-08-459-657-11

Query Match 1.5%; Score 21; DB 2; Length 863;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Qy 1361 TTTGAAAAAAAAAAAAAAAAAAAA 1381

|||||

Db 825 TTTGAAAAAAAAAAAAAAAAAAAA 845

RESULT 38

US-08-460-132-11
; Sequence 11, Application US/08460132
; Patent No. 5928643

GENERAL INFORMATION:

APPLICANT: BIOGEN, INC.
APPLICANT: WALLNER, Barbara P.
APPLICANT: MILLER, Glenn T.
APPLICANT: ROSA, Margaret D.
TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 875 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022-6250
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/459,657
FILING DATE: 02-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02050
FILING DATE: 12-MAR-1992
APPLICATION NUMBER: US 07/667,971
FILING DATE: 12-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,967
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F., Jr.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B151CIP2
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 863 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 18..737
US-08-459-657-11

Query Match 1.5%; Score 21; DB 2; Length 863;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Qy 1361 TTTGAAAAAAAAAAAAAAAAAAAA 1381

|||||

Db 825 TTTGAAAAAAAAAAAAAAAAAAAA 845

NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 875 Third Avenue
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10022-6250
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,132
FILING DATE: 02-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/940,861
FILING DATE: 21-OCT-1992
APPLICATION NUMBER: PCT/US92/02050
FILING DATE: 12-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/667,971
FILING DATE: 12-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,967
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F., Jr.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B151CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 863 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 18..737
US-08-460-132-11

Query Match 1.5%; Score 21; DB 2; Length 863;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTGAAAAA 1381
DB 825 TTGAAAAA 845

RESULT 39
PCT-US92-02050-11
Sequence 11, Application PC/TUS9202050
GENERAL INFORMATION:
APPLICANT: BIOGEN, INC.
APPLICANT: WALLNER, Barbara P.
APPLICANT: MILLER, Glenn T.
APPLICANT: ROSA, Margaret D.
TITLE OF INVENTION: CD2-BINDING DOMAIN OF LYMPHOCYTE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 875 Third Avenue
CITY: New York

STATE: New York
COUNTRY: U.S.A.
ZIP: 10022-6250
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02050
FILING DATE: 19920312
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/667,971
FILING DATE: 12-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,967
FILING DATE: 07-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, James F., Jr.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B151CIP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)715-0600
TELEFAX: (212)715-0673
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 863 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 18..737
PCT-US92-02050-11

Query Match 1.5%; Score 21; DB 5; Length 863;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTGAAAAA 1381
DB 825 TTGAAAAA 845

RESULT 40
5185441-35
Patent No. 5185441
APPLICANT: WALLNER, BARBARA P.; HESSIONS, CATHERINE
TITLE OF INVENTION: DNA SEQUENCES, RECOMBINANT DNA
MOLECULES AND PROCESSES FOR PRODUCING PI-LINKED LYMPHOCYTE
FUNCTION ASSOCIATED ANTIGEN-3
NUMBER OF SEQUENCES: 41
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/237,309
FILING DATE: 26-AUG-1988
SEQ ID NO: 35
LENGTH: 863
5185441-35

Query Match 1.5%; Score 21; DB 6; Length 863;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTGAAAAA 1381
DB 825 TTGAAAAA 845

RESULT 41

```
5223394-5
; Patent No. 5223394
; APPLICANT: WALLNER, BARBARA
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
; LINKAGE SIGNAL SEQUENCES
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/335,688
; FILING DATE: 10-APR-1989
; SEQ ID NO: 5:
; LENGTH: 863
5223394-5
Query Match 1.5%; Score 21; DB 6; Length 863;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTGAAAAA 1381
|||||
DB 825 TTGAAAAA 845

RESULT 42
US-08-616-368A-7
; Sequence 7, Application US/08616368A
; Patent No. 5767262
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Haber, Edgar
; APPLICANT: Jain, Mukesh
; APPLICANT: Yet, Shaw-Fang
; TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/616,368A
; FILING DATE: 15-MAR-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/022001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 880 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; NAME/KEY: Coding Sequence
; LOCATION: 55...633
US-08-616-368A-7
Query Match 1.5%; Score 21; DB 1; Length 880;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTGAAAAA 1381
|||||
```

```
DB 843 TTGAAAAA 863

RESULT 43
US-09-054-298-7
; Sequence 7, Application US/09054298
; Patent No. 6136953
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Haber, Edgar
; APPLICANT: Jain, Mukesh
; APPLICANT: Yet, Shaw-Fang
; TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/054,298
; FILING DATE: 02-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/616,368
; FILING DATE: 15-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Beattie, Ph.D., Ingrid A.
; REGISTRATION NUMBER: P-42,306
; REFERENCE/DOCKET NUMBER: 05433/022002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 880 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; NAME/KEY: Coding Sequence
; LOCATION: 55...633
US-09-054-298-7
Query Match 1.5%; Score 21; DB 3; Length 880;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTGAAAAA 1381
|||||
DB 843 TTGAAAAA 863

RESULT 44
US-08-818-655-7
; Sequence 7, Application US/08818655
; Patent No. 6258557
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; APPLICANT: Haber, Edgar
; APPLICANT: Jain, Mukesh
; APPLICANT: Yet, Shaw-Fang
; TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,655
FILING DATE: 14-MAR-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/616,368
FILING DATE: 15-MAR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/030001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 880 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 55..633
OTHER INFORMATION:
US-08-818-655-7

Query Match 1.5%; Score 21; DB 4; Length 880;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTTGAAAAAAAAAAAAAAAA 1381
DB 843 TTTGAAAAAAAAAAAAAAAA 863

RESULT 45
5223394-8
; Patent No. 5223394
; APPLICANT: WALLNER, BARBARA
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL
; LINKAGE SIGNAL SEQUENCE
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/335,688
; FILING DATE: 10-APR-1989
; SEQ ID NO: 8:
; LENGTH: 1009
5223394-8

Query Match 1.5%; Score 21; DB 6; Length 1009;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1361 TTTGAAAAAAAAAAAAAAAA 1381
DB 973 TTTGAAAAAAAAAAAAAAAA 993

Search completed: November 5, 2002, 13:50:13
Job time : 121.87 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 13:04:56 : Search time 1181.61 Seconds
(without alignments)
15774.429 Million cell updates/sec

Title: US-09-805-311-5

Perfect score: 1381

Sequence: 1 cgaccacgcgtccggccac.....ttgaaaaaaaaaaaaaaaaaa 1381

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estino:*
4: em_estmo:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pin:*
16: em_gss_vit:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	384	27.8	554	10 BE639421	BE639421 946033A02
C 2	371	26.9	470	9 AI861468	AI861468 614014D03
C 3	361	26.1	550	10 BE639422	BE639422 946033A02
C 4	334	24.2	467	9 BE186786	BE186786 946012C08
C 5	317	23.0	553	9 AW562789	AW562789 660065H06
C 6	306	22.2	586	9 AI881599	AI881599 606068G09
C 7	265	19.2	456	9 AI065689	AI065689 ag91f12.x
C 8	263	19.0	901	10 BE837708	BE837708 zm10_01f0
C 9	249	18.0	474	9 AW559173	AW559173 660065H06
C 10	245	17.7	532	9 AW000375	AW000375 614014D03
C 11	241	17.5	475	9 AW562517	AW562517 660065H06
C 12	225	16.3	376	9 AI065546	AI065546 ag8e02.x
C 13	224	16.2	470	9 AW288831	AW288831 707010F11
C 14	192	13.9	414	9 AW288784	AW288784 707010C02
C 15	187	13.5	225	9 AW562788	AW562788 660065H06
C 16	148	10.7	470	9 AW288831	AW288831 707010F11
C 17	130	9.4	363	9 AW562518	AW562518 660065H06

18	103	7.5	232	10	BF727781
C 19	93	6.7	553	9	AI834484
C 20	68	4.9	126	9	AW147048
C 21	65	4.7	126	9	AW147048
C 22	57	4.1	225	9	AI947478
C 23	35	2.5	253	10	D43467
C 24	35	2.5	539	10	BF098320
C 25	35	2.5	592	9	AI780966
C 26	33	2.4	105	9	AI932215
C 27	31	2.2	648	9	AV913663
C 28	30	2.2	220	9	AV414689
C 29	30	2.2	526	10	BI419720
C 30	28	2.0	179	9	AI309589
C 31	27	2.0	278	10	BF713082
C 32	27	2.0	325	10	BF711898
C 33	27	2.0	449	12	CNS020VQ
C 34	27	2.0	460	10	BE417817
C 35	27	2.0	618	10	BI404639
C 36	27	2.0	683	9	AV918966
C 37	27	2.0	754	10	BE416584
C 38	27	2.0	820	10	BG414505
C 39	26	1.9	152	10	BG441951
C 40	26	1.9	212	9	AW089560
C 41	26	1.9	252	10	BG511606
C 42	26	1.9	264	9	AL367841
C 43	26	1.9	276	9	AI349757
C 44	26	1.9	287	9	AI590456
C 45	26	1.9	287	9	AW534167

ALIGNMENTS

RESULT 1
BE639421/C
LOCUS
DEFINITION
ACCESION
VERSION
KEYWORDS
SOURCE
ORGANISM

BE639421 554 bp mRNA linear EST 30-AUG-2000
946033A02.x2 946 - tassal primordium prepared by Schmidt lab Zea
mays cDNA, mRNA sequence.
BE639421
BE639421.1 GI:9952838
EST.
Zea mays.
Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE
AUTHORS
TITLE

Walbot.V.
Maize ESTs from various cDNA libraries sequenced at Stanford

University

Unpublished (1999)

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 946033 row: A column: 02.

Location/Qualifiers

FEATURES

1..554
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassal primordium prepared by Schmidt lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to inflorescence development"
/lab_host="XLOLR"
/note="Organ: tassels; Vector: HybriZAP; Site_1: ECORI; Site_2: XhoI; George Chuck dissected immature tassels between 1mm and 3mm. Sharon Stanfield prepared the cDNA


```

FEATURES source
Location/Qualifiers
1..550
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassal primordium prepared by Schmidt
lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XLOLR"
/note="Organ: tassels; Vector: HybridAP; Site_1: EcoRI;
Site_2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybridAP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."
BASE COUNT 177 a 116 c 149 g 108 t
ORIGIN
Query Match 26.1%; Score 361; DB 10; Length 550;
Best Local Similarity 100.0%; Pred. No. 6.7e-95;
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 186 CATGGAACCTCTCAAAATGAAGCTGTGAAGTCACTAGTCACTTTCGAAGGAATGTTCAA 245
Db 190 CATGGAACCTCTCAAAATGAAGCTGTGAAGTCACTAGTCACTTTCGAAGGAATGTTCAA 249
QY 246 CCGGCAATATAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTTGATGGCAAGCC 305
Db 250 CCGGCAATATAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTTGATGGCAAGCC 309
QY 306 TCTGATATGAAGAAACAAGAGCTTGCTAAAAGATCTCAAAAAGAGATGATGCAACCAA 365
Db 310 TCTGATATGAAGAAACAAGAGCTTGCTAAAAGATCTCAAAAAGAGATGATGCAACCAA 369
QY 366 AGATCTGACTGAGGAGTAGAGGTAGAGATAAAGATGCGATTGAAATTTGAGCAAGAG 425
Db 370 AGATCTGACTGAGGAGTAGAGGTAGAGATAAAGATGCGATTGAAATTTGAGCAAGAG 429
QY 426 GACTCTAAAGGTCAAGGCAACACACGAAGATTTGTAACGCGCTATTAAAGACTTATGGG 485
Db 430 GACTCTAAAGGTCAAGGCAACACACGAAGATTTGTAACGCGCTATTAAAGACTTATGGG 489
QY 486 GGTCTCTGTTGAGAGGCACCTCTGAAGCAGACAGAGATGTCAGCCCTTTGCATAAA 545
Db 490 GGTCTCTGTTGAGAGGCACCTCTGAAGCAGACAGAGATGTCAGCCCTTTGCATAAA 549
QY 546 C 546
Db 550 C 550
RESULT 4
BE186786/c 467 bp mRNA linear EST 22-JUN-2000
LOCUS 946012C08.X1 946 - tassal primordium prepared by Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION BE186786
VERSION BE186786.1 GI:8665970
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 467)
Walbot.V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University

```

```

FEATURES source
Location/Qualifiers
1..467
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassal primordium prepared by Schmidt
lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XLOLR"
/note="Organ: tassels; Vector: HybridAP; Site_1: EcoRI;
Site_2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybridAP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."
BASE COUNT 108 a 110 c 96 g 153 t
ORIGIN
Query Match 24.2%; Score 334; DB 9; Length 467;
Best Local Similarity 99.7%; Pred. No. 4.7e-87;
Matches 384; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 838 AGATATCAAAATTCCTGAGGACTGGCCCTTACCAAGAAGCTCGACGCTTGTTCGAAGGAGCT 897
Db 467 AGATATCAAAATTCCTGAGGACTGGCCCTTACCAAGAAGCTCGACGCTTGTTCGAAGGAGCT 408
QY 898 AATGTCACATTTGGATTATTCCTGAGCTAAATGGACATGACCTGATGAGGAGGCTCTCATA 957
Db 407 AATGTCACATTTGGATTATTCCTGAGCTAAATGGACATGACCTGATGAGGAGGCTCTCATA 348
QY 958 AGTTTCCTGGTAAAGATAATGTTTCAACGAAGATCGGGTGACAAAGGCCATAGAGAAG 1017
Db 347 AGTTTCCTGGTAAAGATAATGTTTCAATGAAGATCGGGTGACAAAGGCCATAGAGAAG 288
QY 1018 ATCAATATCGCAAGAATAATCGTCGAAGGAAGACTCGAGTCCTTTTCAAGCCAAC 1077
Db 287 ATCAATATCGCAAGAATAATCGTCGAAGGAAGACTCGAGTCCTTTTCAAGCCAAC 228
QY 1078 GCCACCATCATGAGCAGCGCTTAAACGGAAGGAGACTTCGGATATAAACAGCAAGCAGCT 1137
Db 227 GCCACCATCATGAGCAGCGCTTAAACGGAAGGAGACTTCGGATATAAACAGCAAGCAGCT 168
QY 1138 GCGAACAAGAAACAAAGGCTGGTGGAAAGAAATAATCTTGGATGCTTGATGTACAA 1197
Db 167 GCGAACAAGAAACAAAGGCTGGTGGAAAGAAATAATCTTGGATGCTTGATGTACAA 108
QY 1198 CTACGACTACGAAGAGCAGCGTGGC 1222
Db 107 CTACGACTACGAAGAGCAGCGTGGC 83
RESULT 5
AW562789 553 bp mRNA linear EST 10-MAR-2000
LOCUS 660065H06.Y1 660 - Mixed stages of anther and pollen Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW562789
VERSION AW562789.1 GI:7216667
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 553)
Walbot.V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University

```

```

JOURNAL      University
COMMENT      Unpublished (1999)
              Contact: Walbot V
              Department of Biological Sciences
              Stanford University
              855 California Ave, Palo Alto, CA 94304, USA
              Tel: 650 723 2227
              Fax: 650 725 8221
              Email: walbot@stanford.edu
              Plate: 660065 row: H column: 06.

FEATURES
  Source      Location/Qualifiers
              1..553
                /organism="Zea mays"
                /cultivar="Ohio43"
                /db_xref="taxon:4577"
                /clone_lib="660 - Mixed stages of anther and pollen"
                /tissue_type="whole premeiotic anthers to pollen shed"
                /dev_stage="premeiotic anthers to pollen shed"
                /lab_host="XLOLR"
                /note="Organ: anthers; Vector: Lambda zap; Site_1: EcoRI;
                Site_2: XhoI; Anther and pollen cDNA library.
                Directionally sequenced with 5' end at the EcoRI site.
                Created by Amie Franklin."
BASE COUNT   171 a 111 c 129 g 142 t
ORIGIN
Query Match   23.0%; Score 317; DB 9; Length 553;
Best Local Similarity 99.7%; Pred. No. 3.3e-82;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 855 GGACTGGCCTTACCAAGAGCTCGACGCTTGTTCGAAGGAGCCTAATGTCACATTGGATAT 914
Db 4 GGACTGGCCTTACCAAGAGCTCGACGCTTGTTCGAAGGAGCCTAATGTCACATTGGATAT 63
Qy 915 TCTGTAGCTAAATGACTGCACCTGATGAGAGGCTCTCATAGTTTCTGTGTTAAAGA 974
Db 64 TCTGTAGCTAAATGACTGCACCTGATGAGAGGCTCTCATAGTTTCTGTGTTAAAGA 123
Qy 975 TAATGTTTTCAAGAGATCGGGTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAA 1034
Db 124 TAATGTTTTCAAGAGATCGGGTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAA 183
Qy 1035 TAATGTCGCAAGAGACTCGACTCTCTTTTCAGCCCACTGCCACACATCAGCACC 1094
Db 184 TAATGTCGCAAGAGACTCGACTCTCTTTTCAGCCCACTGCCACACATCAGCACC 243
Qy 1095 GCTAAACGGAAGGAGACTTCGGATAAACAAGCAAGCGAGCTCGCAACGAAGAAACAAA 1154
Db 244 GCTAAACGGAAGGAGACTTCGGATAAACAAGCAAGCGAGCTCGCAACGAAGAAACAAA 303
Qy 1155 GGCTGGTGAAGAGAAATAATCTTTGGATGCTTGTATGCTACAACTACGACTACGAAGCA 1214
Db 304 GGCTGGTGAAGAGAAATAATCTTTGGATGCTTGTATGCTACAACTACGACTACGAAGCA 363
Qy 1215 GCGGTGTC 1222
Db 364 GCGGTGTC 371

RESULT 6
AI881599      586 bp mRNA linear EST 02-FEB-2000
LOCUS         606068G09.y1 606 - Ear tissue cDNA library from Schmidt lab Zea
DEFINITION   mays cDNA, mRNA sequence.
ACCESSION   AI881599
VERSION     AI881599.1 GI:5566733
KEYWORDS    EST.
SOURCE      Zea mays.
ORGANISM    Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 586)

AUTHORS      Walbot,V.
TITLE        Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL      University
COMMENT      Unpublished (1999)
              Contact: Walbot V
              Department of Biological Sciences
              Stanford University
              855 California Ave, Palo Alto, CA 94304, USA
              Tel: 650 723 2227
              Fax: 650 725 8221
              Email: walbot@stanford.edu
              Plate: 606068 row: G column: 09.

FEATURES
  source      Location/Qualifiers
              1..586
                /organism="Zea mays"
                /cultivar="Ohio43"
                /db_xref="taxon:4577"
                /clone_lib="606 - Ear tissue cDNA library from Schmidt
                lab"
                /tissue_type="mixed"
                /dev_stage="ear length from 0.5 cm - 2.0 cm"
                /lab_host="XLOLR (Stratagene)"
                /note="Organ: immature ear; Vector: pBK-CMV; Site_1: EcoRI
                ; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt
                lab"
BASE COUNT   186 a 132 c 159 g 109 t
ORIGIN
Query Match   22.2%; Score 306; DB 9; Length 586;
Best Local Similarity 99.2%; Pred. No. 4.6e-79;
Matches 506; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 16 GCCACAGCCGCGCAGCAGATGGCATCAAGGGTTTGACGAACTGCTGGGGGCAAT 75
Db 77 GCCACAGCCGCGCAGCAGATGGCATCAAGGGTTTGACGAACTGCTGGGGGCAAT 136
Qy 76 GCGCCCAAGCCATGAAGCAGAGATTCGAGAGCTACTTCGCGCCGCAAAATCGCGTC 135
Db 137 GCGCCCAAGCCATGAAGCAGAGATTCGAGAGCTACTTCGCGCCGCAAAATCGCGTC 196
Qy 136 GAGCCGAGATGAGCATATACCAGTTCTCTGATTTAGTTGGAAGCAGAGCATGGAAC 195
Db 197 GAGCCGAGATGAGCATATACCAGTTCTCTGATTTAGTTGGAAGCAGAGCATGGAAC 256
Qy 196 CTCACAAATGAAGCTGGTCAAGTCACTAGTCTATTTGCAAGGAATGTTCAACCGCAATA 255
Db 257 CTCACAAATGAAGCTGGTCAAGTCACTAGTCTATTTGCAAGGAATGTTCAACCGCAATA 316
Qy 256 AGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGTATGGCAAGCCTCTCTGATG 315
Db 317 AGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGTATGGCAAGCCTCTCTGATG 376
Qy 316 AAGAACAGAGCTTCTTAAAGATACTCAAAAGAGATGATGCAACCAAGATCTGACT 375
Db 377 AAGAACAGAGCTTCTTAAAGATACTCAAAAGAGATGATGCAACCAAGATCTGACT 436
Qy 376 GAGCAGTAGAGGTAGAGATAAAGATCGGATTTGAAAAATTTGACCAAGAGGACTGTAAAG 435
Db 437 GAGCAGTAGAGGTAGAGATAAAGATCGGATTTGAAAAATTTGACCAAGAGGACTGTAAAG 496
Qy 436 GTCACAGGCAACACAAACGAAGTTCTTAAGCGCTATTAAAGACTTATGGGGTTCCTGTT 495
Db 497 GTCACAGGCAACACAAACGAAGTTCTTAAGCGCTATTAAAGACTTATGGGGTTCCTGTT 556
Qy 496 GTAGAGGCACCTTCTTGAAGCAGAGCAAGAA 525
Db 557 GTAGAGGCACCTTCTTGAAGCAGAGCAAGAA 586

RESULT 7
AI065689/c    456 bp mRNA linear EST 24-JUL-1998
LOCUS         ag91f12.x1 maize inflorescence immature ear library Zea mays cDNA
DEFINITION

```

```

clone ag91fl2 3', mRNA sequence.
ACCESSION   AI065689
VERSION     AI065689.1  GI:3341096
KEYWORDS    EST.
SOURCE      Zea mays.
ORGANISM    Zea mays.

REFERENCE   1 (bases 1 to 456)
AUTHORS    Schutz,K., de la Bastide,M., Gnoj,L., Habermann,K., Huang,E.N.,
            Parnell,L.D., Dedhia,N., Martienssen,R. and McCombie,W.R.
TITLE      Expressed sequence tags from Z. mays
JOURNAL    Unpublished (1998)
COMMENT    Contact: W. Richard McCombie
            Lita Annenberg Hazen Genome Sequencing Center
            Cold Spring Harbor Laboratory
            PO Box 100, Cold Spring Harbor, NY 11724, USA
            Tel: 516 367 8884
            Fax: 516 367 8874
            Email: mcombie@cshl.org
            Plate: ag91 row: f column: 12
            Seq primer: M13 forward universal -21
            High quality sequence stop: 456.

FEATURES             source
     location/Qualifiers
     1..456
     /organism="Zea mays"
     /cultivar="B73"
     /db_xref="taxon:4577"
     /clone="ag91fl2"
     /clone_lib="maize inflorescence immature ear library"
     /sex="female"
     /tissue_type="immature ear"
     /note="Vector: pBJUESCRIPT SK+ (X52325); Site_1: XhoI;
     Site_2: EcoRI; This library is described in Schmidt, Hake,
     et al., (1993) Plant Cell 5:729-737. cDNAs are
     directionally cloned into the XhoI and EcoRI sites; XhoI
     is in the polyA tail. Most reads from this library are
     3' in direction. Additional information on this library as
     well as ftp access to all sequences can be found at
     http://www.cshl.org/maizegenome"
BASE COUNT   112 a 104 c 86 g 152 t
ORIGIN
Query Match      19.2%; Score 265; DB 9; Length 456;
Best Local Similarity 100.0%; Pred. No. 3.8e-67;
Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1000 ACAAGGCCATAGAGAGTCAAAATCTGCCAAGATAAATCGTCGCAAGGAGACTCGAG 1059
Db 365 ACAAGGCCATAGAGAGTCAAAATCTGCCAAGATAAATCGTCGCAAGGAGACTCGAG 306

Qy 1060 TCCTTTTCAAGCCAACTGCCACACATCAGCACCGCTTAAACGGAAGGAGACTTCGGAT 1119
Db 305 TCCTTTTCAAGCCAACTGCCACACATCAGCACCGCTTAAACGGAAGGAGACTTCGGAT 246

Qy 1120 AAAACAGCAGGACGCTCGAACAAGAAACAAAGGCTGTGTGAAAGAGAATAATCT 1179
Db 245 AAAACAGCAGGACGCTCGAACAAGAAACAAAGGCTGTGTGAAAGAGAATAATCT 186

Qy 1180 TGGATGCTTGATGTACAACTAGCACTAGCAAGAGCGGTGGCGTGATCACTTCGGTTAG 1239
Db 185 TGGATGCTTGATGTACAACTAGCACTAGCAAGAGCGGTGGCGTGATCACTTCGGTTAG 126

Qy 1240 ATATTATTAACCTCCTGTTTAACTC 1264
Db 125 ATATTATTAACCTCCTGTTTAACTC 101

RESULT 8
LOCUS      BG837708
DEFINITION Zm10_01f08_A Zm10_AAPC_ECORC_Fusarium_graminearum_corn_silk Zea

mays cDNA clone Zm10_01f08, mRNA sequence.
ACCESSION   BG837708
VERSION     BG837708.1  GI:14204031
KEYWORDS    EST.
SOURCE      Zea mays.
ORGANISM    Zea mays.

REFERENCE   1 (bases 1 to 901)
AUTHORS    Harris,L.J., Balcerzak,M., Allard,S., Saparno,A., Couroux,P., De
            Moors,A., Hattori,J.I., Ouellet,T., Robert,L.S., Singh,J.A, Sprott
            , D. and Tinker,N.A.
TITLE      Expressed Sequence Tags from Maize Silk Six Hours After Silk
            Channel Inoculation with Fusarium graminearum
JOURNAL    Unpublished (2001)
COMMENT    Contact: Harris, Linda J.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
            CANADA
            Tel: (613) 759-1314
            Fax: (613) 759-6566
            Email: harris@em.agr.ca.

FEATURES             source
     location/Qualifiers
     1..901
     /organism="Zea mays"
     /cultivar="CO388"
     /db_xref="taxon:4577"
     /clone="Zm10_01f08"
     /clone_lib="Zm10_AAPC_ECORC_Fusarium_graminearum_corn_silk"
     /tissue_type="Silk"
     /dev_stage="4-5 days post-silk emergence"
     /note="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI;
     Site_2: XhoI; Field-grown corn was silk channel-inoculated
     in the morning (~10 am) with 1 ml of a macroconidial
     suspension (500,000 spores/ml) of Fusarium graminearum and
     silk channels were collected and immediately frozen in
     liquid nitrogen 6 hours later. RNA was extracted from
     silk tissue between 1 cm below and above the inoculation
     point in the silk channel, RNA from five silk channels was
     pooled."
BASE COUNT   267 a 174 c 208 g 249 t
ORIGIN
Query Match      19.0%; Score 263; DB 10; Length 901;
Best Local Similarity 99.3%; Pred. No. 8e-67;
Matches 603; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 628 CCAAGTTCACAGAAATACCTCTGATGGAATTTGATGTTGCCAAGTTTGGAGGAGCTT 687
Db 72 CCAAGTTCACAGAAATACCTCTGATGGAATTTGATGTTGCCAAGTTTGGAGGAGCTT 131

Qy 688 GAACCTCACCATGACGACGATTCATTGATTTGTGTCATCTCTGTGATGTCGACTATTGTGAT 747
Db 132 GAACCTCACCATGACGACGATTCATTGATTTGTGTCATCTCTGTGATGTCGACTATTGTGAT 191

Qy 748 AGCATCAAAGGTATCGGGGGCAACACGCTCTGAACCTTATTCGTCACACATGGTCCATA 807
Db 192 AGCATCAAAGGTATCGGGGGCAACACGCTCTGAACCTTATTCGTCACACATGGTCCATA 251

Qy 808 GAAAGCATCTTGGAGAATCTTAAATAAGACAGATATCAAAATTCCTGAGGACTGGCTTA 866
Db 252 GAAAGCATCTTGGAGAATCTTAAATAAGACAGATATCAAAATTCCTGAGGACTGGCTTA 311

Qy 867 CC-AAGAGCTCCAGCTTCTTCAAGGAGCCTTAATGTCACATTCGATATTCCTGACTAA 925
Db 312 CCAAGAGCTCCAGCTTCTTCAAGGAGCCTTAATGTCACATTCGATATTCCTGACTAA 371

Qy 926 AATGGACTGCACCTGATGAGGAGGCTCTCATAAGTTTCTCGTAAAAAGATAATCGTTTCA 985
Db 372 AATGGACTGCACCTGATGAGGAGGCTCTCATAAGTTTCTCGTAAAAAGATAATCGTTTCA 431

```



```

QY 1025 CTGCCAAGTAATTCGTCGACGAGACTCGAGTCCTTTTCACGCCAATGCCACCA 1084
|||||
Db 408 CTGCCAAGTAATTCGTCGACGAGACTCGAGTCCTTTTCACGCCAATGCCACCA 467
|||||
QY 1085 CATCAGCCCGCTAAACGGAGAGAGACTTCGGATAAAACAAAGCAGCGTGGGAACA 1144
|||||
Db 468 CATCAGCCCGCTAAACGGAGAGAGACTTCGGATAAAACAAAGCAGCGTGGGAACA 527
|||||
QY 1145 AGAAA 1149
|||||
Db 528 AGAAA 532
|||||

RESULT 11
AW562517/c
LOCUS
DEFINITION
60065H06.X2 660 - Mixed stages of anther and pollen Zea mays cDNA,
mRNA sequence.
ACCESSION AW562517
VERSION AW562517.1 GI:7216395
KEYWORDS
EST.
SOURCE
Zea mays.
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 475)
AUTHORS
Walbot,V.
TITLE
Maize ESTs from various cDNA libraries sequenced at Stanford
University
JOURNAL
Unpublished (1999)
COMMENT
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 60065 row: H column: 06.
Location/Qualifiers
1..475
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XL0LR"
/Note="Organ: anthers; Vector: Lambda zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 124 a 110 c 94 g 147 t
ORIGIN
Query Match 17.5%; Score 241; DB 9; Length 475;
Best Local Similarity 99.7%; Pred. No. 3.1e-60;
Matches 291; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 931 ACTGCACCTGATGAGAGGGTCTCATAGTTTCTCGTAAAGATATATGTTTCACGAA 990
|||||
Db 475 ACTGCACCTGATGAGAGGGTCTCATAGTTTCTCGTAAAGATATATGTTTCACGAA 416
|||||
QY 991 GATCGGGTGACAAAGCCATAGAAAGATCAAAATCTCCCAAGATTAATCGTCGCAAGA 1050
|||||
Db 415 GATCGGGTGACAAAGCCATAGAAAGATCAAAATCTCCCAAGATTAATCGTCGCAAGA 356
|||||
QY 1051 AGACTCGAGTCCTTTTTCAGCGCACTGCCACACATCAGCACCGGTAAACGGAGGAG 1110
|||||
Db 355 AGACTCGAGTCCTTTTTCAGCGCACTGCCACACATCAGCACCGGTAAACGGAGGAG 296
|||||
QY 1111 ACTTCGGATAAACCAAGCAGCGACTCGGAACAAGAAACCAAGGCTGGTGAAGAAG 1170
|||||

```

```

Db 295 ACTTCGGATAAAACAAGCAGCGAGCTGCCAACAAAGAACAAAGCGCTGGTGAAGAAG 236
|||||
QY 1171 AAATAATCTTGATGCTTGATGTACAACTACGACTACGAAAGCAGCGGTGGC 1222
|||||
Db 235 AAATAATCTTGATGCTTGATGTACAACTACGACTACGAAAGCAGCGGTGGC 184
|||||

RESULT 12
AI065546/c
LOCUS
DEFINITION
AI065546
ag88e02.x1 maize inflorescence immature ear library Zea mays cDNA
clone ag88e02 3', mRNA sequence.
ACCESSION AI065546
VERSION AI065546.1 GI:3340953
KEYWORDS
EST.
SOURCE
Zea mays.
ORGANISM
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 376)
AUTHORS
Schutz,K., de la Bastide,M., Gnoj,L., Habermann,K., Huang,E.N.,
Parnell,L.D., Dedhia,N., Martienssen,R. and McCombie,W.R.
TITLE
Expressed sequence tags from Z. mays
JOURNAL
Unpublished (1998)
COMMENT
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: ag88 row: e column: 02
Seq primer: M13 forward universal -21
High quality sequence stop: 376.
Location/Qualifiers
1..376
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ag88e02"
/clone_lib="maize inflorescence immature ear library"
/sex="female"
/tissue_type="immature ear"
/Note="Vector: pBLUESCRIPT SK+ (X52325); Site_1: XhoI;
Site_2: EcoRI; This library is described in Schmidt, Hake,
et al., (1993) Plant Cell 5:729-737. cDNAs are
directionally cloned into the XhoI and EcoRI sites; XhoI
is near the polyA tail. Most reads from this library are
3' in direction. Additional information on this library as
well as ftp access to all sequences can be found at
http://www.cshl.org/maizegenome"
BASE COUNT 79 a 99 c 65 g 131 t
ORIGIN
Query Match 16.3%; Score 225; DB 9; Length 376;
Best Local Similarity 99.6%; Pred. No. 1.5e-55;
Matches 275; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 968 TAAAGATAATCGTTTCAACGAGATCGGGTGACAAAGCCATAGAGAGATCAATCTG 1027
|||||
Db 376 TAAAGATAATCGTTTCAACGAGATCGGGTGACAAAGCCATAGAGAGATCAATCTG 317
|||||
QY 1028 CCAAGATAAATCGTCGCAAGGAGACTCGAGTCCTTTTCAAGCCCACTGCCACCACAT 1087
|||||
Db 316 CCAAGATAAATCGTCGCAAGGAGACTCGAGTCCTTTTCAAGCCCACTGCCACCACAT 257
|||||
QY 1088 CAGCACCCTGTAACGAGGAGAGACTTCGGATAAAACAAAGCAGCGTGGCAACAAGA 1147
|||||
Db 256 CAGCACCCTGTAACGAGGAGAGACTTCGGATAAAACAAAGCAGCGTGGCAACAAGA 197
|||||
QY 1148 AAACAAGGCTGGTGGAAAGAGAAATAATCTTGATGCTTGTACAACTACGACTAC 1207
|||||

```

```

Db 196 AACAAGGCTGGTGAAGAAGAAATAATCTTGATGCTGATGATACAGACTAC 137
QY 1208 GAACGACGGTGGCGTGATCACTTCGCTTAGATTA 1243
Db 136 GAACGACGGTGGCGTGATCACTTCGCTTAGATTA 101

RESULT 13
AW288831/c
LOCUS
DEFINITION
470 bp mRNA linear EST 16-JAN-2000
mays cDNA, mRNA sequence.
ACCESSION
AW288831
VERSION
AW288831.1 GI:5695753
KEYWORDS
EST.
SOURCE
Zea mays.
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 470)
AUTHORS
Walbot,V.
TITLE
Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL
University
COMMENT
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707010 row: F column: 11.
FEATURES
source
1..470
Location/Qualifiers
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK
)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Organ: tassel, kernel, silk, husk, root, leaf;
vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."
BASE COUNT
102 a 110 c 108 g 150 t
ORIGIN
Query Match 16.2%; Score 224; DB 9; Length 470;
Best Local Similarity 99.6%; Pred. No. 2.5e-35;
Matches 274; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 914 TTCTTGAGCTAAATGGAGTCTGACCTGATGAGGAGGCTCTCATAAGTTTCCTGGTAAAG 973
Db 429 TTCTTGAGCTAAATGGAGTCTGACCTGATGAGGAGGCTCTCATAAGTTTCCTGGTAAAG 370
QY 974 ATAAATGTTTCAACGAAGATCGGTGACAAAGGCCATAGAGAAGATCAAACTGCGCAAGA 1033
Db 369 ATAAATGTTTCAACGAAGATCGGTGACAAAGGCCATAGAGAAGATCAAACTGCGCAAGA 310
QY 1034 ATAAATCGTCGCAAGGAAGATCGAGTCTCTTTTCAAGCCAACTGCCACACATCAGCAC 1093
Db 309 ATAAATCGTCGCAAGGAAGATCGAGTCTCTTTTCAAGCCAACTGCCACACATCAGCAC 250
QY 1094 CGCTAAACGGAGAGACTTCGATATAAACAACAGGAGGAGCTGGCAACAGAAACAA 1153
Db 249 CGCTAAACGGAGAGACTTCGATATAAACAACAGGAGGAGCTGGCAACAGAAACAA 190
QY 1154 AGCTGGTGAAGAAGAAATAATCTTGGATGCTT 1188
|||||

```

```

Db 189 AGCTGGTGAAGAAGAAATAATCTTGGATGCTT 155

RESULT 14
AW288784
LOCUS
DEFINITION
414 bp mRNA linear EST 16-JAN-2000
mays cDNA, mRNA sequence.
ACCESSION
AW288784
VERSION
AW288784.1 GI:5695706
KEYWORDS
EST.
SOURCE
Zea mays.
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 414)
AUTHORS
Walbot,V.
TITLE
Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL
University
COMMENT
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707010 row: C column: 02.
FEATURES
source
1..414
Location/Qualifiers
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK
)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Organ: tassel, kernel, silk, husk, root, leaf;
vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."
BASE COUNT
120 a 82 c 101 g 109 t 2 others
ORIGIN
Query Match 13.9%; Score 192; DB 9; Length 414;
Best Local Similarity 99.0%; Pred. No. 4.6e-46;
Matches 392; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 601 CCACGGTTCCTTCGTCATTAAATGATCCCAAGTTCCCAAGAAATACCTGTGATGGAATT 660
Db 6 CCACGGTTCCTTCGTCATTAAATGATCCCAAGTTCCCAAGAAATACCTGTGATGGAATT 65
QY 661 GATGTTGCCAAGGTTTGGAGAGCTTGAACCTCACCATGGACCAAGTTCATTGATTTGTC 720
Db 66 GATGTTGCCAAGGTTTGGAGAGCTTGAACCTCACCATGGACCAAGTTCATTGATTTGTC 125
QY 721 ATCTGTGTGGATGTGACTATTGTGATAGCATCAAGGTATCGGGGGGCAACAGCTCTG 780
Db 126 ATCTGTGTGGATGTGACTATTGTGATAGCATCAAGGTATCGGGGGGCAACAGCTCTG 185
QY 781 AAATCTATTTCGTCACATCGGTCCATAGAAAGCATCTTGGAGATCTTTAATAAGACAGA 840
Db 186 AAATCTATTTCGTCACATCGGTCCATAGAAAGCATCTTGGAGATCTTTAATAAGACAGA 245
QY 841 TATCAATTCCTGAGACTGGCCTTACCAGAAAGCTGCAGCTTGTTCAGAGGACCTAAT 900
Db 246 TACCAATTCCTGAGACTGGCCTTACCAGAAAGCTGCAGCTTGTTCAGAGGACCTAAT 305
QY 901 GTCACATGGAATATTCCTGAGCTAAATGGACTGCACCTGATGAGGAGGCTCTCATAAGT 960
|||||

```


TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
 JOURNAL University
 COMMENT Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 660065 row: H column: 06.

FEATURES
 source
 1. .363
 /organism="Zea mays"
 /cultivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="660 - Mixed stages of anther and pollen"
 /tissue_type="whole premiotic anthers to pollen shed"
 /dev_stage="premiotic anthers to pollen shed"
 /lab_host="XLOLR"
 /note="Organ: Anthers; Vector: Lambda zap; Site_1: EcoRI;
 Site_2: XhoI; Anther and pollen cDNA library.
 Directionally sequenced with 5' end at the EcoRI site.
 Created by Amie Franklin."
 82 c 74 g 111 t

BASE COUNT 96 a 82 c 74 g 111 t
 ORIGIN

Query Match 9.4%; Score 130; DB 9; Length 363;
 Best Local Similarity 100.0%; Pred. No. 3.9e-28;
 Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1093 CCCTAAACGGAAGCAGACTTCGGATAAAACAAGCAGCGTGGCAACAAAGAAACA 1152
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 301 CCCTAAACGGAAGCAGACTTCGGATAAAACAAGCAGCGTGGCAACAAAGAAACA 242
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1153 AAGCTGGTGGGAAGAGAAATATCTTGGATGCTTGATGATACACTACGACTACGAAG 1212
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 241 AAGCTGGTGGGAAGAGAAATATCTTGGATGCTTGATGATACACTACGACTACGAAG 182
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 1213 CAGCGGTGGC 1222
 |||||||||

DB 181 CAGCGGTGGC 172

RESULT 18
 BF727781 232 bp mRNA linear EST 09-JAN-2001
 LOCUS
 DEFINITION
 CDNA, mRNA sequence.
 ACCESSION BF727781
 VERSION BF727781.1 GI:12045642
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 232)
 AUTHORS Walbot V.
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
 JOURNAL University
 COMMENT Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 100052 row: F column: 12.

FEATURES
 source
 1. .232
 /organism="Zea mays"

/db_xref="dbEST:707010C02.y1"
 /db_xref="taxon:4577"
 /clone_lib="1000 - Unigene I from Maize Genome Project"
 /note="This library represents the unique ESTs found in
 the first round of EST sequencing at Stanford University
 for the maize genome project. Sequences are present from
 libraries 486, 487, 496, 603, 605, 606, 614, 618, 660, 683
 , 687, 707, and 945. Contigs were assembled using TIGR's
 CAP program and a representative EST from each contig was
 selected for the Unigene set. All singlets were also
 selected."

BASE COUNT 64 a 42 c 57 g 69 t
 ORIGIN

Query Match 7.5%; Score 103; DB 10; Length 232;
 Best Local Similarity 100.0%; Pred. No. 3.5e-20;
 Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 707 TCATTGATTGTCATCCTGCTGCTGCTGACTATTTGATAGCATCAAGGTATCGGG 766
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 55 TCATTGATTGTCATCCTGCTGCTGCTGACTATTTGATAGCATCAAGGTATCGGG 114
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 767 GCGAACAGCTCTGAAACATATTCGTCAACATGGTCCATAGA 809
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 115 GCGAACAGCTCTGAAACATATTCGTCAACATGGTCCATAGA 157
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 19
 AI834484/c
 LOCUS
 DEFINITION
 606068G09.xl 606 - Ear tissue cDNA library from Schmidt lab Zea
 mays cDNA, mRNA sequence.
 ACCESSION AI834484
 VERSION AI834484.1 GI:5468693
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 553)
 AUTHORS Walbot V.
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
 JOURNAL University
 COMMENT Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 606068 row: G column: 09.

FEATURES
 source
 1. .553
 /organism="Zea mays"
 /cultivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="606 - Ear tissue cDNA library from Schmidt
 lab"
 /tissue_type="mixed"
 /dev_stage="ear length from 0.5 cm - 2.0 cm"
 /lab_host="XLOLR (Stratagene)"
 /note="Organ: immature ear; Vector: pBK-CMV; Site_1: EcoRI
 ; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt
 lab"

BASE COUNT 140 a 112 c 106 g 195 t
 ORIGIN

Query Match 6.7%; Score 93; DB 9; Length 553;
 Best Local Similarity 100.0%; Pred. No. 1.3e-17;
 Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 988 GAAGTCGGTGACAAAGCCCATAGACAGATCAAAATCTGCCAAGAATAATCTGCCAA 1047
|||||
Db 512 CAAGATCGGGTGACAAAGCCCATAGAGAGATCAAAATCTGCCAAGAATAATCTGCCAA 453
|||||
QY 1048 GGAAGACTCGAGTCTCTTTTCAAGCCCAACTGCC 1080
|||||
Db 452 GGAAGACTCGAGTCTCTTTTCAAGCCCAACTGCC 420
|||||

RESULT 20
AW147048
LOCUS AW147048 126 bp mRNA linear EST 03-NOV-1999
DEFINITION 707010C02.y1 707 - Mixed adult tissues from Walbot lab (SK) Zea
mays cDNA, mRNA sequence.
ACCESSION AW147048
VERSION AW147048.1 GI:6194944
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 126)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707010 row: C column: 02.

FEATURES
source
location/Qualifiers
1..126
/organism="Zea mays"
/db_xref="W23"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)
)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

BASE COUNT 29 a 31 c 26 g 40 t
ORIGIN
1..126
/organism="Zea mays"
/db_xref="W23"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)
)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

Query Match 4.9%; Score 68; DB 9; Length 126;
Best Local Similarity 100.0%; Pred. No. 7.3e-10;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 749 GCATCAAGGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCAACATGGGTCATAG 808
|||||
Db 1 GCATCAAGGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCAACATGGGTCATAG 60
|||||

BASE COUNT 29 a 31 c 26 g 40 t
ORIGIN
1..126
/organism="Zea mays"
/db_xref="W23"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)
)"
/tissue_type="tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

Query Match 4.9%; Score 68; DB 9; Length 126;
Best Local Similarity 100.0%; Pred. No. 7.3e-10;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 809 AAAGCATC 816
|||||
Db 61 AAAGCATC 68
|||||

RESULT 21
AW147048/c
LOCUS AW147048 126 bp mRNA linear EST 03-NOV-1999
DEFINITION 707010C02.y1 707 - Mixed adult tissues from Walbot lab (SK) Zea
mays cDNA, mRNA sequence.
ACCESSION AW147048
VERSION AW147048.1 GI:6194944

```

```

KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 126)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707010 row: C column: 02.

FEATURES
source
location/Qualifiers
1..126
/organism="Zea mays"
/db_xref="W23"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)
)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

BASE COUNT 29 a 31 c 26 g 40 t
ORIGIN
1..126
/organism="Zea mays"
/db_xref="W23"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)
)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

Query Match 4.7%; Score 65; DB 9; Length 126;
Best Local Similarity 100.0%; Pred. No. 5.4e-09;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1124 CAAGCAAGCGAGCTGCCAACAAGAAAACAAAGCTGTGTGGAAGAAGAAATAATCTTGA 1183
|||||
Db 126 CAAGCAAGCGAGCTGCCAACAAGAAAACAAAGCTGTGTGGAAGAAGAAATAATCTTGA 67
|||||

QY 1184 TGCTT 1188
|||||
Db 66 TGCTT 62
|||||

RESULT 22
AW147478/c
LOCUS AW147478 225 bp mRNA linear EST 19-AUG-1999
DEFINITION 614047B01.x1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
mRNA sequence.
ACCESSION AW147478
VERSION AW147478.1 GI:5739683
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 225)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL University
COMMENT Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227

```

```

Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614047 row: B column: 01.
Location/Qualifiers
1. .225
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="614 - root cDNA library from Walbot Lab"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XL0LR"
/notes=Organ: root; Vector: pBluescriptII SK-; site_1:
EcORI; site_2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM)
BASE COUNT      66 a 54 c 40 g 65 t
ORIGIN
Query Match      4.1%; Score 57; DB 9; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.6e-07;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1137 TCGCAACAGAAACAAAGCGTGTGGAAAGAAATAATCTTGGATGCTTGATGT 1193
|||||
DB 225 TCGCAACAGAAACAAAGCGTGTGGAAAGAAATAATCTTGGATGCTTGATGT 169
|||||

RESULT 23
D43467
LOCUS      253 bp mRNA linear EST 04-MAY-1998
DEFINITION D43467 Rice callus cDNA (H.Uchimiya) Oryza sativa cDNA clone S8447,
mRNA sequence.
ACCESSION  D43467
VERSION    D43467.1 GI:3107727
KEYWORDS   EST.
SOURCE     Oryza sativa.
ORGANISM   Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 253)
Uchimiya,H.
On nucleotide sequence of Oryza sativa
Unpublished (1994)
Contact: Hirofumi Uchimiya
Institute of Mol. & Cell. Bioscience, Department of Cellular
Function
The University of Tokyo
1-1-1 Yayoi, Bunkyo-ku, Tokyo 113, Japan
Tel: 03-3812-2111(ex.7844)
Fax: 03-3812-2910
Email: huchimiya@tansei.cc.u-tokyo.ac.jp
PROJECT= Uchimiya.
Location/Qualifiers
1. .253
/organism="Oryza sativa"
/db_xref="taxon:4530"
/clone="SS447"
/clone_lib="Rice callus cDNA (H.Uchimiya)"
/tissue_type="callus"
BASE COUNT      72 a 42 c 64 g 74 t 1 others
ORIGIN
Query Match      2.5%; Score 35; DB 10; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 469 CTATTAGACTTATGGGGTTCCTGTTGTAGAGCC 503
|||||
DB 34 CTATTAGACTTATGGGGTTCCTGTTGTAGAGCC 68
|||||

RESULT 24
D43467
LOCUS      592 bp mRNA linear EST 18-MAY-2001
DEFINITION D43467 Rice callus cDNA (H.Uchimiya) Oryza sativa cDNA clone S8447,
mRNA sequence.
ACCESSION  D43467
VERSION    D43467.1 GI:5279007
KEYWORDS   EST.
SOURCE     tomato.
ORGANISM   Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 592)
D'Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,

```

TITLE
JOURNAL
COMMENT

Giovannoni, J.J. and Martin, G.B.
Generation of ESTs from *Pseudomonas* susceptible tomato
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES

source

1. .592
/organism="Lycopersicon esculentum"
/cultivar="R11-13 (Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="cLES13P2"
/clone_lib="tomato susceptible, Cornell"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLES - Tomato *Pseudomonas* Susceptible EST Library.
Directionally cloned cDNAs inserted into pBlueScript SK(-)
at 5' end with EcoRI and 3' end with XhoI site"

BASE COUNT
ORIGIN

190 a 109 c 149 g 144 t

Query Match 2.5%; Score 35; DB 9; Length 592;
Best Local Similarity 100.0%; Pred. No. 0.64;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 GGAATCAAGCCAGTTATGTTTGTGATGGCAAGCC 305

|||||
Db 306 GGAATCAAGCCAGTTATGTTTGTGATGGCAAGCC 340

RESULT 26
AI932215/C

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

618029H11.x1 618 - Inbred Tassel cdNA Library Zea mays cDNA, mRNA
sequence.
AI932215
AI932215.1 GI:5670952
EST.
Zea mays.
Zea mays.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.

Walbot.V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V

Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 618029 row: H column: 11.

FEATURES
source

1. .105
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="618 - Inbred Tassel cdNA Library"
/tissue_type="tassel"
/dev_stage="tassel length from 0.1 to 2.5 cm"
/lab_host="XLOLR"
/note="Organ: tassel; Vector: pAD-GAL4-2.1 (Hybrizap);
Inbred tassel library from Schmidt lab"

BASE COUNT

36 a 27 c 21 g 21 t

ORIGIN

Query Match 2.4%; Score 33; DB 9; Length 105;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1284 GCCCATGTTCAAGCTGGGTAAAGTTAGTTGTG 1316
|||||
Db 33 GCCCATGTTCAAGCTGGGTAAAGTTAGTTGTG 1

RESULT 27

AV913663

LOCUS

AV913663 K. sato unpublished cdNA library, cv. Haruna Niho
germination shoots Hordeum vulgare subsp. vulgare cdNA clone
bags22n17 5', mRNA sequence.

ACCESSION

AV913663

VERSION

AV913663.1

GI:18209440

KEYWORDS

EST.

SOURCE

Hordeum vulgare subsp. vulgare.

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.

REFERENCE

1 (bases 1 to 648)

AUTHORS

Sato, K., Saisho, D. and Takeda, K.

TITLE

Barley EST sequencing project in NIG and Okayama Univ

JOURNAL

Unpublished (2002)

COMMENT

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tsnig@genes.nig.ac.jp.

FEATURES

source

1. .648

/organism="Hordeum vulgare subsp. vulgare"

/cultivar="Haruna Niho"

/db_xref="taxon:112509"

/clone="bags22n17"

/clone_lib="K. Sato unpublished cdNA library, cv. Haruna

Niho germination shoots"

/tissue_type="shoots"

/dev_stage="germination"

BASE COUNT 174 a 158 c 185 g 131 t

ORIGIN

Query Match 2.2%; Score 31; DB 9; Length 648;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 GGAGCAGAGTTCGAGAGCTACTTCGCGCCG 123
|||||
Db 138 GGAGCAGAGTTCGAGAGCTACTTCGCGCCG 168

RESULT 28

AV414689

LOCUS

DEFINITION

AV414689 Lotus japonicus young plants (two-week old) Lotus

japonicus cdNA clone MM247H08_r 5', mRNA sequence.

ACCESSION

AV414689

VERSION

AV414689.1

GI:7743865

KEYWORDS

EST.

SOURCE

Lotus japonicus.

ORGANISM

Lotus japonicus.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
Lotus.

REFERENCE

1 (bases 1 to 220)

AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE Generation of 7137 non-redundant expressed sequence tags from a
 legume, *Lotus japonicus*
JOURNAL DNA Res. 7 (2), 127-130 (2000)
MEDLINE 20277479
COMMENT Contact: Yasukazu Nakamura
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: ynakamu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES Location/Qualifiers
 1..220

/organism="Lotus japonicus"

/db_xref="taxon:34305"

/clone="MM247h08_r"

/clone_lib="Lotus japonicus young plants (two-week old)"

/dev_stage="young plants (two-week old)"

/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:

XhoI; Isolate-Miyakojima MG-20"

54 a 66 c 35 g 65 t

BASE COUNT 54 a 66 c 35 g 65 t
ORIGIN

Query Match 2.2%; Score 30; DB 9; Length 220;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTCGAGAGCTACTTCGGCGCAAAATCGCC 132

|||||

DB 158 TTCGAGAGCTACTTCGGCGCAAAATCGCC 187

RESULT 29

LOCUS BI419720

DEFINITION LJNES7479r Lotus japonicus nodule library 5 and 7 week-old Lotus japonicus cDNA 5', mRNA sequence.
ACCESSION BI419720
VERSION BI419720.1
KEYWORDS EST.
SOURCE Lotus japonicus.

ORGANISM Lotus japonicus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids 1; Fabales; Fabaceae; Papilionoideae; Loteae;

Lotus.

1 (bases 1 to 526)

Colbatch, G., Freund, S., Trevaskis, B and Udvardi, M.

Lotus japonicus root nodule ESTs: tools for functional genomics

Unpublished (2000)

Contact: Udvardi, MK

Molecular Plant Nutrition

Max Planck Institute of Molecular Plant Physiology

An Muehlenberg 1, 14476 Golm, Germany

Fax: 49 331 567 8250

Email: udvardi@mpimp-golm.mpg.de

Seq primer: 17

High quality sequence stop: 526.

FEATURES Location/Qualifiers

1..526

/organism="Lotus japonicus"

/cultivar="Gifu (B-129)"

/db_xref="taxon:34305"

/clone_lib="Lotus japonicus nodule library 5 and 7

week-old"

/dev_stage="5 and 7 week-old plants"

/note="Organ: Nodule; Vector: pSPORT1; Site_1: SalI;

Site_2: NotI. The library was prepared using mRNA

extracted from nodules of 5 and 7 week-old Lotus plants.

Nodules were induced by, and contained Mesorhizobium

strain R7A."

170 a 113 c 127 g 116 t

BASE COUNT 170 a 113 c 127 g 116 t
ORIGIN

Query Match 2.2%; Score 30; DB 10; Length 526;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTCGAGAGCTACTTCGGCGCAAAATCGCC 132

|||||

DB 109 TTCGAGAGCTACTTCGGCGCAAAATCGCC 138

RESULT 30

LOCUS AI309589

DEFINITION tb29c02.x1 NCI-CGAP Kid12 Homo sapiens cDNA clone IMAGE:2055746 3', similar to contains element MER22 repetitive element ;, mRNA sequence.
ACCESSION AI309589
VERSION AI309589.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 179)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 911 Std Error: 0.00

Seq primer: -400P from Gibco

High quality sequence stop: 173.

FEATURES Location/Qualifiers

1..179

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NCI-CGAP Kid12"

/tissue_type="2 pooled tumors (clear cell type)"

/lab_host="DH10B"

/note="Organ: kidney; Vector: p7T3D-pac (Pharmacia) with

a modified polylinker; Site_1: Not I; Site_2: Eco RI;

Plasmid DNA from the normalized library NCI-CGAP Kid5 was

prepared, and ss circles were made in vitro. Following HAP

purification, this DNA was used as tracer in a subtractive

hybridization reaction. The driver was PCR-amplified cDNAs

from a pool of 5,000 clones made from the same library

(cloneIDs 132912-1325831, 1471368-1472903 and

1492104-1493255). Subtraction by Bento Soares and M.

Fatima Bonaldo."

54 a 18 c 22 g 85 t

BASE COUNT 54 a 18 c 22 g 85 t
ORIGIN

Query Match 2.0%; Score 28; DB 9; Length 179;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1354 GCTGTTTGTGAAAAAATAAAAAA 1381

|||||

DB 147 GCTGTTTGTGAAAAAATAAAAAA 174

RESULT 31
LOCUS BF713082

DEFINITION 278 bp mRNA linear EST 02-JAN-2001

```

DEFINITION MI-P-H3-adm-f-10-1-UM.s1 MI-P-H3 Sus scrofa cDNA clone
ACCESSION MI-P-H3-adm-f-10-1-UM 3', mRNA sequence.
VERSION BF713082
KEYWORDS BF713082.1 GI:12012557
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 278)
AUTHORS Bonaudo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
COMMENT Contact: Tugle CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kilgus Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401
Email: cktugle@iastate.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
non-normalized hypothalamus at estrus day 12 library cDNA library
Preparation: RJ Woods, JA Green, RS Prather S142 Animal Science
Research Center, Department of Animal Science, University of
Missouri-Columbia, 65211 Clone distribution: clones will be
available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes
Location/Qualifiers
1..278
/organism="Sus scrofa"
/strain="crossbred"
/db_xref="taxon:9823"
/clone="MI-P-H3-adm-f-10-1-UM"
/clone_lib="MI-P-H3"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: EcoRI; The MI-P-H3
library is derived from hypothalamus at estrus day 12.
For a detailed description of the library from which this
clone was derived, please visit our web site at
http://pigest.genome.iastate.edu/."
TAG_LIB="MI-P-H3"
TAG_TISSUE="hypothalamus at estrus day 12"
TAG_SEQ="GGTAA"
BASE COUNT 83 a 46 c 44 g 105 t
ORIGIN
Query Match 2.0%; Score 27; DB 10; Length 278;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1355 CTGTTTTTGAAGAAAAAAGAAAAA 1381
|||||
Db 28 CTGTTTTTGAAGAAAAAAGAAAAA 2

RESULT 32
BF711898/c
LOCUS BF711898
DEFINITION MI-P-O3-abg-d-08-1-UM.s1 MI-P-O3 Sus scrofa cDNA clone
ACCESSION BF711898
VERSION BF711898.1 GI:12011373
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 278)
AUTHORS Bonaudo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
COMMENT Contact: Tugle CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kilgus Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401
Email: cktugle@iastate.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
non-normalized hypothalamus at estrus day 12 library cDNA library
Preparation: RJ Woods, JA Green, RS Prather S142 Animal Science
Research Center, Department of Animal Science, University of
Missouri-Columbia, 65211 Clone distribution: clones will be
available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes
Location/Qualifiers
1..278
/organism="Sus scrofa"
/strain="crossbred"
/db_xref="taxon:9823"
/clone="MI-P-H3-adm-f-10-1-UM"
/clone_lib="MI-P-H3"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: EcoRI; The MI-P-H3
library is derived from hypothalamus at estrus day 12.
For a detailed description of the library from which this
clone was derived, please visit our web site at
http://pigest.genome.iastate.edu/."
TAG_LIB="MI-P-H3"
TAG_TISSUE="hypothalamus at estrus day 12"
TAG_SEQ="GGTAA"
BASE COUNT 83 a 46 c 44 g 105 t
ORIGIN
Query Match 2.0%; Score 27; DB 10; Length 278;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1355 CTGTTTTTGAAGAAAAAAGAAAAA 1381
|||||
Db 28 CTGTTTTTGAAGAAAAAAGAAAAA 2

RESULT 32
BF711898/c
LOCUS BF711898
DEFINITION MI-P-O3-abg-d-08-1-UM.s1 MI-P-O3 Sus scrofa cDNA clone
ACCESSION BF711898
VERSION BF711898.1 GI:12011373
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 278)
AUTHORS Bonaudo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
COMMENT Contact: Tugle CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kilgus Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401
Email: cktugle@iastate.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
non-normalized hypothalamus at estrus day 12 library cDNA library
Preparation: RJ Woods, JA Green, RS Prather S142 Animal Science
Research Center, Department of Animal Science, University of
Missouri-Columbia, 65211 Clone distribution: clones will be
available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes
Location/Qualifiers
1..325
/organism="Sus scrofa"
/strain="crossbred"
/db_xref="taxon:9823"
/clone="MI-P-O3-abg-d-08-1-UM"
/clone_lib="MI-P-O3"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: EcoRI; The MI-P-O3
library is derived from ovary at estrus day 12. For a
detailed description of the library from which this clone
was derived, please visit our web site at
http://pigest.genome.iastate.edu/."
TAG_LIB="MI-P-O3"
TAG_TISSUE="ovary at estrus day 12"
TAG_SEQ="TTGTAC"
BASE COUNT 104 a 56 g 115 t
ORIGIN
Query Match 2.0%; Score 27; DB 10; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1355 CTGTTTTTGAAGAAAAAAGAAAAA 1381
|||||
Db 28 CTGTTTTTGAAGAAAAAAGAAAAA 2

RESULT 33
CNS020YQ/c
LOCUS CNS020YQ
DEFINITION Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
224018 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL176219
VERSION AL176219.1 GI:7814276
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 449)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fitzames,C., Fisher,C.,

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 325)
AUTHORS Bonaudo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
COMMENT Contact: Tugle CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kilgus Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401
Email: cktugle@iastate.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to verify it as a clone from the
non-normalized ovary at estrus day 12 library cDNA library
Preparation: RJ Woods, JA Green, RS Prather S142 Animal Science
Research Center, Department of Animal Science, University of
Missouri-Columbia, 65211 Clone distribution: clones will be
available through Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=Yes
Location/Qualifiers
1..325
/organism="Sus scrofa"
/strain="crossbred"
/db_xref="taxon:9823"
/clone="MI-P-O3-abg-d-08-1-UM"
/clone_lib="MI-P-O3"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: EcoRI; The MI-P-O3
library is derived from ovary at estrus day 12. For a
detailed description of the library from which this clone
was derived, please visit our web site at
http://pigest.genome.iastate.edu/."
TAG_LIB="MI-P-O3"
TAG_TISSUE="ovary at estrus day 12"
TAG_SEQ="TTGTAC"
BASE COUNT 104 a 56 g 115 t
ORIGIN
Query Match 2.0%; Score 27; DB 10; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1355 CTGTTTTTGAAGAAAAAAGAAAAA 1381
|||||
Db 28 CTGTTTTTGAAGAAAAAAGAAAAA 2

RESULT 33
CNS020YQ/c
LOCUS CNS020YQ
DEFINITION Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
224018 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL176219
VERSION AL176219.1 GI:7814276
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 449)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fitzames,C., Fisher,C.,

```

Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis

JOURNAL REFERENCE

2 (bases 1 to 449)
 Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence

JOURNAL REFERENCE

3 (bases 1, to 449)
 Direct Submission
 Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases

JOURNAL COMMENT

This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

source
 1..449
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="224018"
 /note="Genoscope sequence ID : COAG224BH09SP1-end ; PUC-ori"

BASE COUNT 70 a 72 c 48 g 215 t 44 others
 ORIGIN

Query Match 2.0%; Score 27; DB 12; Length 449;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1355 CACGTTTGTGAAAAAAGAAAAA 1381
 |||||

Db 424 CAGTTTTTGAAAAAAGAAAAA 398

RESULT 34
 BE417817
 LOCUS
 DEFINITION
 MUG024.H12R90620 ITEC MUG wheat Spikelet Library Triticum aestivum
 CDNA clone MUG024.H12, mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE

BE417817
 BE417817.1 GI:9415663
 EST.
 bread wheat.
 ORGANISM
 Triticum aestivum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
 ; Triticeae; Triticum.

REFERENCE

1 (bases 1 to 460)
 Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier,S., Dubcovsky,J., Feuillet,C., Gale,M., Graner,A., Gustafson,P., Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P., Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y., Pecchioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M., Sorrells,M., Warburton,M. and Wenzel,G.
 International Triticeae EST Cooperative (ITEC): Production of Expressed Sequence Tags for Species of the Triticeae
 Unpublished (2000)
 Contact: Ogihara Y
 Maifara Institute for Biological Research, Yokohama City University
 Kohoka-cho 641-12, Totsuka-ku, Yokohama 244-0813, JAPAN
 Tel: 81 45 820 1903
 Fax: 81 45 820 1901
 Email: ogihara@yokohama-cu.ac.jp
 International Triticeae EST Cooperative (ITEC)
<http://wheat.pw.usda.gov/genome>.
 Location/Qualifiers
 1..460

FEATURES

source
 1..460
 /organism="Triticum aestivum"
 /cultivar="Norin 26"
 /db_xref="taxon:4565"
 /clone="MUG024.H12"
 /clone_lib="ITEC MUG Wheat Spikelet Library"
 /tissue_type="young spikelets"
 /dev_stage="Feekes' scale 6-7"
 /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; M13 Reverse sequencing primer used. 1.2 kbp average insert size."

BASE COUNT 142 a 89 c 112 g 117 t
 ORIGIN

Query Match 2.0%; Score 27; DB 10; Length 460;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 840 ATATCAAAATTCCTGAGGACTGGCCTTA 866
 |||||

Db 187 ATATCAAAATTCCTGAGGACTGGCCTTA 213

RESULT 35
 BE404639/c
 LOCUS
 DEFINITION
 MI-P-NA-aea-g-05-1-UM.s1.abl MI-P-NA Sus scrofa cDNA clone
 MI-P-NA-aea-g-05-1-UM.abl 3', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE

BE404639.1 GI:15183852
 EST.
 pig.
 ORGANISM
 Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.

REFERENCE

1 (bases 1 to 616)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene discovery
 Genome Res. 6 (9), 791-806 (1996)
 9704477

CONTACT: Tuggle CK
 Molecular Genetics Laboratory, Department of Animal Science
 Iowa State University
 201 Kilgus Hall, Ames, IA 50011-3150, USA
 Tel: 5152944252
 Fax: 5152942401
 Email: ckting@iastate.edu

The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized anterior pituitary at estrus day 12 library cDNA Library Preparation: RJ Woods, JA Green, RS Prather S142 Animal Science Research Center, Department of Animal Science, University of Missouri-Columbia, 65211 Clone Distribution: clones will be available through Research Genetics (www.resgen.com)
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES

Location/Qualifiers
 1..618
 /organism="Sus scrofa"
 /strain="crossbreed"
 /db_xref="taxon:9823"
 /clone="MI-P-NA-aea-g-05-1-UM.abl"
 /clone_lib="MI-P-NA"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: NotI; Site_2: EcoRI; The MI-P-NA is a normalized library comprised of anterior pituitary tissue at estrus days 0, 5, and 12. For a detailed description of the library from which this clone was derived, please visit our web site at <http://pigest.genome.iastate.edu/>.


```

TAG_LIB=MI-P-NA
TAG_TISSUE=anterior pituitary at estrus day 12
TAG_SEQ=TCACAG"
BASE COUNT      174 a 116 c 109 g 219 t
ORIGIN

Query Match      2.0%; Score 27; DB 10; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1355 CCGTTTTTGAAGAAAAAAGAAAAA 1381
      |||
Db 28 CCGTTTTTGAAGAAAAAAGAAAAA 2

RESULT 36
AV918966/c
LOCUS
DEFINITION
AV918966 683 bp mRNA linear EST 18-JAN-2002
AV918966 K. Sato unpublished cDNA library, cv. Haruna Nijo
germination shoots Hordeum vulgare subsp. vulgare cDNA clone
bags22n17 3', mRNA sequence.
AV918966
AV918966.1 GI:18214745
EST.
SOURCE
Hordeum vulgare subsp. vulgare.
ORGANISM
Hordeum vulgare subsp. vulgare

REFERENCE
AUTHORS Sato,K., Saisho,D. and Takeda,K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. 683
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
/clone="bags22n17"
/tissue="bags22n17"
/dev_stage="shoots"
/germination="shoots"
/germination="shoots"

BASE COUNT      175 a 169 c 134 g 203 t
ORIGIN

Query Match      2.0%; Score 27; DB 9; Length 683;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 840 ATATCAAAATTCCTGAGGACTGCGCTTA 866
      |||
Db 529 ATATCAAAATTCCTGAGGACTGCGCTTA 503

RESULT 37
BE416584
LOCUS
DEFINITION
MUG009.B10R990122 ITEC MUG wheat Spikelet Library Triticum aestivum
cDNA clone MUG009.B10, mRNA sequence.
BE416584
AV918966.1 GI:9414332
EST.
SOURCE
bread wheat.
Triticum aestivum
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;


```

```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Triticum.
1 (bases 1 to 754)
Anderson,O.A., Appels,R., Bailey,P., Blake,T., Close,T., Cloutier
,S., Dubcovsky,J., Feuillet,C., Gale,M., Granet,A., Gustafson,P.,
Herrmann,R.G., Holton,T., Jacquemin,J.M., Jia,J., Joudrier,P.,
Langridge,P., Lazo,G.R., Lin,J.J., McGuire,P., Ogihara,Y.,
Pechioni,N., Qualset,C., Schuch,W., Selvaraj,G., Shariflou,M.,
Sorrells,M., Warburton,M. and Wenzel,G.
International Triticale EST Cooperative (ITEC): Production of
Expressed Sequence Tags for Species of the Triticaceae
Unpublished (2000)
Contact: Ogihara Y
Kihara Institute for Biological Research, Yokohama City University
Maio-cho 641-12, Totsuka-ku, Yokohama 244-0813, JAPAN
Tel: 81 45 820 1903
Fax: 81 45 820 1901
Email: ogihara@yokohama-cu.ac.jp
International Triticale EST Cooperative (ITEC)
http://wheat.pw.usda.gov/genome.
Location/Qualifiers
1. 754
/organism="Triticum aestivum"
/cultivar="Norin 26"
/db_xref="taxon:4565"
/clone="MUG009.B10"
/clone_lib="ITEC MUG Wheat Spikelet Library"
/tissue_type="young spikelets"
/dev_stage="Feekes' scale 6-7"
/Note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; M13 Reverse sequencing primer used. 1.2 kbp average
insert size."

BASE COUNT      225 a 149 c 181 g 199 t
ORIGIN

Query Match      2.0%; Score 27; DB 10; Length 754;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 840 ATATCAAAATTCCTGAGGACTGCGCTTA 866
      |||
Db 187 ATATCAAAATTCCTGAGGACTGCGCTTA 213

RESULT 38
BG414505
LOCUS
DEFINITION
HVSMEK002L10f Hordeum vulgare testa/pericarp EST library
HVCBNA0013 (normal) Hordeum vulgare cDNA clone HVSMEK002L10f, mRNA
sequence.
BG414505
BG414505.1 GI:13320056
EST.
SOURCE
Hordeum vulgare
barley.
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 820)
Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Kannangara,G., von
Wetstein,D., Ahniov,E., Chin,A., Cho,D.W., Panton,R.D., Kianian
,P., Otto,C., Simons,K., Zhang,D., Begum,D., Frisch,D., Yu,X.,
Henry,D., Palmer,M., Rambo,T., Simmons,J., Oates,R. and Main,D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex testa/pericarp cDNA library
Unpublished (2001)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu


```

Total hg bases = 427
Seq primer: AATTACCCCTCACTAAGGG
High quality sequence stop: 744.
Location/Qualifiers

FEATURES

source

1..820
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEK0002L10f"
/clone_lib="Hordeum vulgare testa/pericarp EST library
HVCDA0013 (normal)"
/tissue_type="testa/pericarp"
/lab_host="TJC121"
/note="vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
Plants were raised from seeds in a Controlled Environments
growth chamber maintained in continuous light at 18°C, and
testa and pericarp were dissected from developing kernels
at Washington State University, Pullman, WA (Kannaugara,
von Wettstein). Total RNA was prepared, poly(A) RNA was
purified, one cDNA library was made, and 1 million pfu
phagemids in the TJ Clonase lab at the University of
California, Riverside (Akhunov, Chin, Choi, Close, Fenton,
Kianian, Otto, Simons, Zhang). Phagemids were plated and
picked at the Clemson University Genomics Institute (CUGI)
(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
preparations, DNA sequencing and sequence analysis were
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates
, Rambo, Main). The sequence has been trimmed to remove
vector sequence and contains a minimum of 100 bases of
phred value 20 or above. For more details on library
preparation and sequence analysis see
<http://www.genome.clemson.edu/projects/barley>. To order
this clone see <http://www.genome.clemson.edu/orders> Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT 258 a 165 c 190 g 207 t
ORIGIN
Query Match 2.0%; Score 27; DB 10; Length 820;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 840 ATATCAAAATCTCGAGACTGGCCTTA 866
|||||
Db 187 ATATCAAAATCTCGAGACTGGCCTTA 213
|||||

RESULT 39

BG441951
LOCUS
DEFINITION
GA_Ea0015D21f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboreum cDNA clone GA_Ea0015D21f, mRNA sequence.
BAC441951
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Gossypium arboreum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 152)
Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
D., Wood,T.C., Leslie,A. and Wilkins,J.A.
An integrated analysis of the genetics, development, and evolution
of the cotton fiber
Unpublished (2000)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
High quality sequence start: 5
High quality sequence stop: 152.
Location/Qualifiers

FEATURES

source

1..152
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0015D21f"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 52 a 15 c 28 g 57 t
ORIGIN

Query Match 1.9%; Score 26; DB 10; Length 152;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1356 TGTGTTTTGAAAAAATAAAAAA 1381
|||||
Db 126 TGTGTTTTGAAAAAATAAAAAA 151
|||||

RESULT 40

AW089560/c
LOCUS
DEFINITION
xd20c04.xl NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2594310 3',
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 212)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/hcicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor gene index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -400p from Gibco.

FEATURES

source

1..212
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2594310"
/clone_lib="NCI_CGAP_Ov23"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.35 Kb. tumor types include: mixed
Mullerian tumor, papillary serous, clear cell, spindle
cell. All are primary tumors, metastasis positive. Life
Technologies catalog #: 11534-013"
BASE COUNT 60 a 37 c 37 g 78 t

Qy 1356 TGTTTTTGAAGAAAAAAGAAAAA 1381
 |||||||
 Db 238 TGTTTTTGAAGAAAAAAGAAAAA 263
 |||||||

RESULT 43
 AI349757
 LOCUS
 DEFINITION ta96c11.x1 NCI_CGAP_Lu26 Homo sapiens cDNA clone IMAGE:2051924 3', mRNA linear EST 16-FEB-1999
 mRNA sequence.

ACCESSION AI349757
 VERSION AI349757
 KEYWORDS
 SOURCE EST.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 276)

AUTHORS NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 cDNA Library Preparation: David B. Krizman, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
 Insert Length: 154 Std Error: 0.00
 Seq primer: -40UP from Gibco.

FEATURES
 source
 1..276
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2051924"
 /clone_lib="NCI_CGAP_Lu26"
 /tissue_type="invasive adenocarcinoma"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: lung; Vector: pAMP1; mRNA made from lung adenocarcinoma tissue, cDNA made by oligo-dT priming. Directionally cloned. Size-selected on agarose gel, average insert size 500 bp. Primary library, non-amplified." 80 g 54 t
 BASE COUNT 100 a 42 c 80 g 54 t
 ORIGIN

Query Match 1.9%; Score 26; DB 9; Length 276;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1356 TGTTTTTGAAGAAAAAAGAAAAA 1381
 |||||||
 Db 226 TGTTTTTGAAGAAAAAAGAAAAA 251
 |||||||

RESULT 44
 AI590456
 LOCUS
 DEFINITION tt76h07.x1 NCI_CGAP_HSC3 Homo sapiens cDNA clone IMAGE:2246749 3', mRNA linear EST 21-APR-1999
 mRNA sequence.

ACCESSION AI590456
 VERSION AI590456
 KEYWORDS
 SOURCE EST.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 287)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Herbert Morse, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: David B. Krizman, Ph.D.
 DNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Gibco.

FEATURES
 source
 1..287
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2246749"
 /clone_lib="NCI_CGAP_HSC3"
 /tissue_type="CD34+, T negative, patient with chronic myelogenous leukemia"
 /lab_host="DH10B"
 /note="Organ: bone marrow; Vector: pAMP1; mRNA made from lymphoid tissue, cDNA made by oligo-dT priming. Directionally cloned. Size-selected on agarose gel, average insert size 500 bp. Primary library, non-amplified. cDNA Library Preparation: David B. Krizman, Ph.D. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383." 81 g 54 t
 BASE COUNT 106 a 46 c 81 g 54 t
 ORIGIN

Query Match 1.9%; Score 26; DB 9; Length 287;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1356 TGTTTTTGAAGAAAAAAGAAAAA 1381
 |||||||
 Db 237 TGTTTTTGAAGAAAAAAGAAAAA 262
 |||||||

RESULT 45
 AW534167
 LOCUS
 DEFINITION UI-R-C4-aiq-a-03-0-UI.s1 UI-R-C4 Rattus norvegicus cDNA clone UI-R-C4-aiq-a-03-0-DI 3', mRNA linear EST 06-MAR-2000
 ACCESSION AW534167
 VERSION AW534167
 KEYWORDS
 SOURCE EST.
 ORGANISM Norway rat

REFERENCE Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 287)
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 9704477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu

Oligo-dT track not found, Not 1 site shown in beginning of sequence is likely internal to the message. cDNA library preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com)
 Seq primer: M13 Forward

POLYA-NO. Location/Qualifiers
1. .287
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C4-alq-a-03-0-UI"
/clone_lib="UI-R-C4"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="vector: pT73D-pac (Pharmacia) with a modified
polylinker; Site_1: Not 1; Site_2: Eco RI; The UI-R-C4
library is a subtracted library of a series, ultimately
derived from a mixture of tissues from rat placenta,
adult lung, brain, liver, kidney, heart, spleen, ovary,
muscle, and 8, 12 and 18-day embryos. For a detailed
description of the library from which this clone was
derived, please visit our web site at
ratest.eng.uiowa.edu. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)
TAG_LIB=UI-R-C4
TAG_TISSUE=ovary
TAG_SEQ=TCAC"
BASE COUNT 64 a 72 c 72 g 79 t
ORIGIN
Query Match 1.9%; Score 26; DB 9; Length 287;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1356 TGTGTTTTGAAAAAAGAAAAA 1381
|||||
Db 251 TGTGTTTTGAAAAAAGAAAAA 276

Search completed: November 5, 2002, 16:22:17
Job time : 1210.61 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 5, 2002, 11:03:23 : Search time 1874.34 Seconds
(without alignments)
16501.458 Million cell updates/sec

Title: US-09-805-311-7
Perfect score: 1478
Sequence: 1 cgaaccacgcgtccgggaaa.....aaaaaaaaaaaaaaaaaaaaa 1478

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
------------	-------	-------	--------	-------	-------------

1	1478	100.0	1478	6	ARI52406	ARI52406 Sequence
2	894	60.5	1463	6	ARI52403	ARI52403 Sequence
3	850	57.5	1381	6	ARI52405	ARI52405 Sequence
4	842	57.0	1541	6	ARI52404	ARI52404 Sequence
5	58	3.9	1354	8	AB021666	AB021666 Oryza sat
6	58	3.9	115907	2	AC104713	AC104713 Oryza sat
7	32	2.2	178914	9	AL590807	AL590807 Human DNA
8	30	2.0	851	6	AX147252	AX147252 Sequence
9	30	2.0	1487	9	AK027130	AK027130 Homo sapi
10	30	2.0	2168	9	BC020218	BC020218 Homo sapi
11	30	2.0	2899	9	BSM000643	BSM000643 Homo sapi
12	30	2.0	110811	8	AP004520	AP004520 Lotus jap
13	30	2.0	133139	2	AC103556	AC103556 Rattus no
14	30	2.0	150164	10	AF342999	AF342999 Mus muscu
15	30	2.0	162201	2	AC024152	AC024152 Homo sapi
16	30	2.0	191134	9	AC006065	AC006065 Homo sapi
17	30	2.0	208360	9	AC020708	AC020708 Homo sapi
18	30	2.0	233048	2	AC096327	AC096327 Rattus no
19	30	2.0	244849	2	AC090887	AC090887 Mus muscu
20	29	2.0	372	9	BC015361	BC015361 Homo sapi
21	29	2.0	871	8	CAR010224	AJ010224 Cicer ari
22	29	2.0	1018	6	AX022319	AX022319 Sequence
23	29	2.0	1018	6	AX030786	AX030786 Sequence
24	29	2.0	1018	6	BD007470	BD007470 Remedies
25	29	2.0	1029	6	AR048809	AR048809 Sequence
26	29	2.0	1029	6	AR048810	AR048810 Sequence
27	29	2.0	1029	6	AR050404	AR050404 Sequence
28	29	2.0	1029	6	AR050405	AR050405 Sequence
29	29	2.0	1143	9	HUMORFK	L40401 Homo sapien
30	29	2.0	1166	9	BC014413	BC014413 Homo sapi
31	29	2.0	1243	9	BSM802621	AL162083 Homo sapi
32	29	2.0	1449	9	BC011460	BC011460 Homo sapi
33	29	2.0	1488	10	AF244361	AF244361 Mus muscu
34	29	2.0	1601	3	AY061540	AY061540 Drosophil
35	29	2.0	1642	10	BC014817	BC014817 Mus muscu
36	29	2.0	1644	10	AF327431	AF327431 Mus muscu
37	29	2.0	1665	9	BC002488	BC002488 Homo sapi
38	29	2.0	1692	9	AB050426	AB050426 Macaca fa
39	29	2.0	1710	10	BC003901	BC003901 Mus muscu
40	29	2.0	1902	9	AK021579	AK021579 Homo sapi
41	29	2.0	1997	9	BC009569	BC009569 Homo sapi
42	29	2.0	2146	9	AF006083	AF006083 Homo sapi
43	29	2.0	2542	5	XRNAFAS	X75938 X.laevvis mr
44	29	2.0	2804	10	S69407	S69407 Edg-endogli
45	29	2.0	2937	9	BC008857	BC008857 Homo sapi

ALIGNMENTS

RESULT 1	ARI52406	1478 bp	DNA	linear	PAT 08-AUG-2001
LOCUS	Sequence 7 from patent US 6232527.				
DEFINITION	ARI52406				
ACCESSION	ARI52406				
VERSION	ARI52406.1	GI:15118456			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1. (bases 1 to 1478)				
AUTHORS	Mahajan, P.B.				
TITLE	Maize Rad2/FEN-1 orthologues and uses thereof				
JOURNAL	Patent: US 6232527-A 7 15-MAY-2001;				
FEATURES	Location/Qualifiers				
source	1. 1478				
BASE COUNT	463 a 302 c 365 g 348 t				
ORIGIN	/organism="unknown"				

Query Match 100.0%; Score 1478; DB 6; Length 1478;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CGACCCACGGTCCGGAAATAGCTCGCGGTTCGGCTTCTTTCGGCCACTCCGGCTCAGC 60
Db 1 CGACCCACGGTCCGGAAATAGCTCGCGGTTCGGCTTCTTTCGGCCACTCCGGCTCAGC 60
QY 61 CGCGCCCGCCACCGCCACAGCCCGCCACAGAGATGGCATCAAGGGTTTGGAGAAA 120
Db 61 CGCGCCCGCCACCGCCACAGCCCGCCACAGAGATGGCATCAAGGGTTTGGAGAAA 120
QY 121 CTGCTGGCGCAATTCGCGCCCAAGCGATGAAGGAGCAGAGTTTCGAGACTACTTCGCG 180
Db 121 CTGCTGGCGCAATTCGCGCCCAAGCGATGAAGGAGCAGAGTTTCGAGACTACTTCGCG 180
QY 181 CGCAAAATCGCGTGGACGCCAGCATGAGCATCTACAGTTTCCTGATAGTGTGGAAGG 240
Db 181 CGCAAAATCGCGTGGACGCCAGCATGAGCATCTACAGTTTCCTGATAGTGTGGAAGG 240
QY 241 ACAGGCATGGAACCTCTCACAATGAAGCTGGTGAAGTCACTAGTCAATTTGCAAGGAATG 300
Db 241 ACAGGCATGGAACCTCTCACAATGAAGCTGGTGAAGTCACTAGTCAATTTGCAAGGAATG 300
QY 301 TTCAACCGCACAAATAGATTACTGAAGCGGAATCAAGCCAGTTTATGTTTGTGATGGC 360
Db 301 TTCAACCGCACAAATAGATTACTGAAGCGGAATCAAGCCAGTTTATGTTTGTGATGGC 360
QY 361 AAGCCTCCTGATATGAAGAAACAAGAACTTCTTAAAGATACTCAAAAGAGATGATGCA 420
Db 361 AAGCCTCCTGATATGAAGAAACAAGAACTTCTTAAAGATACTCAAAAGAGATGATGCA 420
QY 421 ACCAAAGATCTGACTGAGGCAGTAGAGTAGAGATGAAGATGGGATGAAAATTTGAGC 480
Db 421 ACCAAAGATCTGACTGAGGCAGTAGAGTAGAGATGAAGATGGGATGAAAATTTGAGC 480
QY 481 AAGAGGACTGTAAAGGTCAAGGCAACACAAACGAAGTTCTAAACGACTATTTAAGACIT 540
Db 481 AAGAGGACTGTAAAGGTCAAGGCAACACAAACGAAGTTGTAAACGACTATTTAAGACIT 540
QY 541 ATGGGGTTCCTGTGTAGAGCAGCTTCTGAAGCAGAGCAAGATGTGAGCCCTTTGC 600
Db 541 ATGGGGTTCCTGTGTAGAGCAGCTTCTGAAGCAGAGCAAGATGTGAGCCCTTTGC 600
QY 601 ATAAAGATAAGGTGTCTGCTTCTTTCAGAGATAGGACTCCCTTACTTTTGGGGCT 660
Db 601 ATAAAGATAAGGTGTCTGCTTCTTTCAGAGATAGGACTCCCTTACTTTTGGGGCT 660
QY 661 CCACGGTTCCTTGTGATTTAATGATCCAAAGTTTCCAAAGAAATACCTGTGATGAATTT 720
Db 661 CCACGGTTCCTTGTGATTTAATGATCCAAAGTTTCCAAAGAAATACCTGTGATGAATTT 720
QY 721 GATGTTGCCANGTTTGGAGAGCTTGAACCTCACCATGGACAGTTTCATTTGATTTGTC 780
Db 721 GATGTTGCCANGTTTGGAGAGCTTGAACCTCACCATGGACAGTTTCATTTGATTTGTC 780
QY 781 ATCCGTGTGGATGTGACTATTGTATAGCATCAAGGTATCGGGGGCAACAGCTCTG 840
Db 781 ATCCGTGTGGATGTGACTATTGTATAGCATCAAGGTATCGGGGGCAACAGCTCTG 840
QY 841 AAACCTTATTCGTCACATGGGTCCATAGAAGCATCTTGGAGATCTTAATAAGACAGA 900
Db 841 AAACCTTATTCGTCACATGGGTCCATAGAAGCATCTTGGAGATCTTAATAAGACAGA 900
QY 901 TATCAAAATTCCTGAGACTTGGCTTTACCAAGAAGCTCGACCTTGTTCAGAGAGCCTAAT 960
Db 901 TATCAAAATTCCTGAGACTTGGCTTTACCAAGAAGCTCGACCTTGTTCAGAGAGCCTAAT 960
QY 961 GTCACCTTGGATATTCCTGAGCTAAAATGAGCTGACCTGATAGAGGGTCTCATAAGT 1020
Db 961 GTCACCTTGGATATTCCTGAGCTAAAATGAGCTGACCTGATAGAGGGTCTCATAAGT 1020
QY 1021 TTCTGTAAAGATAATCGTTTCAATGAAGATCGGGTGACAAAGGCCATAGAGAGATC 1080
Db 1021 TTCTGTAAAGATAATCGTTTCAATGAAGATCGGGTGACAAAGGCCATAGAGAGATC 1080
```

```
QY 1081 AATCTGCCAAGATAAATCGTCSAAGSAGACTCGAGTCTTTTTCAGGCCAACTGCC 1140
Db 1081 AATCTGCCAAGATAAATCGTCSAAGSAGACTCGAGTCTTTTTCAGGCCAACTGCC 1140
QY 1141 ACCACATCAGCAGCCGCTAAACGGAAGGAGACTTCGATAAACAAGCAGGAGCTGG 1200
Db 1141 ACCACATCAGCAGCCGCTAAACGGAAGGAGACTTCGATAAACAAGCAGGAGCTGG 1200
QY 1201 AACAGAAACAAGGCTGGTGGAAAGAAATAATCTTGGATGCTTGATGTACAATA 1260
Db 1201 AACAGAAACAAGGCTGGTGGAAAGAAATAATCTTGGATGCTTGATGTACAATA 1260
QY 1261 CGACTACGAAAGCAGCGGTGGCATGATCAGCTTCGCTAGATTATTAACTCCCTGTTTA 1320
Db 1261 CGACTACGAAAGCAGCGGTGGCATGATCAGCTTCGCTAGATTATTAACTCCCTGTTTA 1320
QY 1321 ACTCAGACCTTTGGTGAAGTTTGGCCATGTTCAAGCTGGGGTAAGTTAGTTGTTTG 1380
Db 1321 ACTCAGACCTTTGGTGAAGTTTGGCCATGTTCAAGCTGGGGTAAGTTAGTTGTTTG 1380
QY 1381 AAGAGATTGCTGTACCAAGTAACAAAACCTTATCGCTGTTTTTTTACTTCTTCTGCTTGA 1440
Db 1381 AAGAGATTGCTGTACCAAGTAACAAAACCTTATCGCTGTTTTTTTACTTCTTCTGCTTGA 1440
QY 1441 GTATGTATGCCAGTAAAAAAGAAAAAAGAAAAA 1478
Db 1441 GTATGTATGCCAGTAAAAAAGAAAAAAGAAAAA 1478

RESULT 2
AR152403
LOCUS AR152403 1463 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6232527.
ACCESSION AR152403
VERSION AR152403.1 GI:15118453
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1463)
AUTHORS Mahajan, P.B.
TITLE Maize Rad2/FEN-1 orthologues and uses thereof
JOURNAL Patent: US 6232527-A, 1 15-MAY-2001;
FEATURES Location/Qualifiers
source
1..1463
/organism="unknown"
BASE COUNT 466 a 292 c 361 g 344 t
ORIGIN

Query Match 60.5%; Score 894; DB 6; Length 1463;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1244; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 32 CGCGGTTTCTTCGCGCCACTCCGGCTCAGCGCGCGCCGCCACCGCCACAGCGCCGCGAG 91
Db 20 CGCGGTTTCTTCGCGCCACTCCGGCTCAGCGCGCGCGCCACCGCCACAGCGCCGCGAG 79
QY 92 ACAGATCGGCATCAAGGTTTGACGAAACTCTGGCGGACAATCGCGCCCAAGCGCATGA 151
Db 80 ACAGATCGGCATCAAGGTTTGACGAAACTCTGGCGGACAATCGCGCCCAAGCGCATGA 139
QY 152 AGGAGCAGAAGTTCGAGAGCTACTTCGGCGCCGCAAAATCGCGCTCGAGCCGAGCATGACA 211
Db 140 AGGAGCAGAAGTTCGAGAGCTACTTCGGCGCCGCAAAATCGCGCTCGAGCCGAGCATGACA 199
QY 212 TCTACCAAGTTCTCTGATAGTAGTTGAGAGCAGAGCATGGAACCTCTCACAATAAGACTG 271
Db 200 TATACCAAGTTCTCTGATAGTAGTTGAGAGCAGAGCATGGAACCTCTCACAATAAGACTG 259
QY 272 GTGAAGTCACTAGTCAATTTTCAAGGAATGTTCAACCGGACAAATAGATTACTGGAAGCGG 331
Db 260 GTGAAGTCACTAGTCAATTTTCAAGGAATGTTCAACCGGACAAATAGATTACTGGAAGCGG 319
```


Qy	332	GAATCAAGCCAGTTTATGTTTTGATGGCAAGCCTCTCGATATGAAGAAACAAGAACTTGG	391
Db	320	GAATCAAGCCAGTTTATGTTTTGATGGCAAGCCTCTCGATATGAAGAAACAAGAACTTGG	379
Qy	392	CTAAAAGATACTCAAAAAGAGATGATGCAACAAAGATCTGACTGAGGCAGTAGAGGTAG	451
Db	380	CTAAAAGATACTCAAAAAGAGATGATGCAACCAAAAGATCTGACTGAGGCAGTAGAGGTAG	439
Qy	452	GAGATAAAGATCGGATTGAAAATTTGACAGAGGAGCTGTAAAGGTTCACAAGGCCAACACA	511
Db	440	GAGATAAAGATCGGATTGAAAATTTGACAGAGGAGCTGTAAAGGTTCACAAGGCCAACACA	499
Qy	512	ACGAAGATTGTTAAAGCACTATTAAGACTTATGGGGTTCCTGTTGTAGAGGCACCTTCTG	571
Db	500	ACGAAGATTGTTAAAGCGCTATTAGACTTATGGGGTTCCTGTTGTAGAGGCACCTTCTG	559
Qy	572	AAGCAGAAGCAAGATGTCAGCCCTTTGCCATAACGATAGGTTCCTCGCTGTTGCTTCAG	631
Db	560	AAGCAGAAGCAAGATGTCAGCCCTTTGCCATAACGATAGGTTCCTCGCTGTTGCTTCAG	619
Qy	632	AAGATAAGGACTCCCTTTACTTTTGGGGCTCCACGGTTCCTTCCTCATTTTAATGGATCCAA	691
Db	620	AAGATAAGGACTCCCTTTACTTTTGGGGCTCCACGGTTCCTTCCTCATTTTAATGGATCCAA	679
Qy	692	GTTCGAAGAAATACCTGTGATGGAAATTTGATTTGCCAAGTTTTCGAGGAGCTTGAAC	751
Db	680	GTTCGAAGAAATACCTGTGATGGAAATTTGATTTGCCAAGTTTTCGAGGAGCTTGAAC	739
Qy	752	TCACCATGGACCAAGTTCANTGATTTGTGCATCCTGTGTGGATGTGACTATTGTGTATGACA	811
Db	740	TCACCATGGACCAAGTTCANTGATTTGTGCATCCTGTGTGGATGTGACTATTGTGTATGACA	799
Qy	812	TCAAAGGTATCGGGGGGCAACAGCTCTGAACTTATTCGTCAACATGGGTCCATAGAA	871
Db	800	TCAAAGGTATCGGGGGGCAACAGCTCTGAACTTATTCGTCAACATGGGTCCATAGAA	859
Qy	872	GCATCTTGGAGAACTCTTAATAAGACAGATATCAAAATTCCTGAGGAGTCGCCCTTACCAAG	931
Db	860	GCATCTTGGAGAACTCTTAATAAGACAGATATCAAAATTCCTGAGGAGTCGCCCTTACCAAG	919
Qy	932	AAGCTCGAGCCTTGTTCAGGAGGCTTAATGTCACTTTTGGATATTCCTGAGCTAAAATGGA	991
Db	920	AAGCTCGAGCCTTGTTCAGGAGGCTTAATGTCACTTTTGGATATTCCTGAGCTAAAATGGA	979
Qy	992	CTGCACCTGATGAGGAGGCTCTCATTAAGTTTCTGGTAAAGATATGTTTCAATGAAG	1051
Db	980	CTGCACCTGATGAGGAGGCTCTCATTAAGTTTCTGGTAAAGATATGTTTCAACGAAG	1039
Qy	1052	ATCGGGTGCACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGATAAATCGTCGCAAGGA	1111
Db	1040	ATCGGGTGCACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGATAAATCGTCGCAAGGA	1099
Qy	1112	GACTCGAGTCTTTTTCAGGCCAATGCCACACATCAGCACCGCTTAACGGAAGGAGA	1171
Db	1100	GACTCGAGTCTTTTTCAGGCCAATGCCACACATCAGCACCGCTTAACGGAAGGAGA	1159
Qy	1172	CTTCGGATAAACACAGCAAGGCAGCTGCGAACAAGAAACAAAGGCTTGGTGGAAAGAAGA	1231
Db	1160	CTTCGGATAAACACAGCAAGGCAGCTGCGAACAAGAAACAAAGGCTTGGTGGAAAGAAGA	1219
Qy	1232	AATAATCTTGGATGCTTGTATGTACAACTTACGACTACGAAAGACGCGGTGGC	1282
Db	1220	AATAATCTTGGATGCTTGTATGTACAACTTACGACTACGAAAGACGCGGTGGC	1270
RESULT 3			
AR152405			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
Unknown.			
<div> <div>AR152405</div> <div>Sequence 5 from patent US 6232527.</div> <div>AR152405.1</div> <div>GI:15118455</div> </div>			
<div> <div>1381 bp</div> <div>DNA</div> <div>linear</div> <div>PAT 08-AUG-2001</div> </div>			

```
Db 856 GACTGGCCTTACCAAGAAGCTCGACGCTTGTTCAGGAGCCTAAATCTCACATTCGATATT 915
Qy 976 CTTGAGCTTAAATGGAGTGCACCTGATGAGAGGGTCTCATAGTTTCTGGTAAAGAT 1035
Db 916 CTTGAGCTTAAATGGAGTGCACCTGATGAGAGGGTCTCATAGTTTCTGGTAAAGAT 975
Qy 1036 AATGTTTCAATGAAGATCGSGTGCACAAAGCCCATAGAGAAGATCAAAATCTGCCAAGAT 1095
Db 976 AATGTTTCAACGAAGATCGSGTGCACAAAGCCCATAGAGAAGATCAAAATCTGCCAAGAT 1035
Qy 1096 AATCGTCGCAAGAGACTCGAGTCTTTTCAAGCCAACTGCCACCATCAGCACCG 1155
Db 1036 AATCGTCGCAAGAGACTCGAGTCTTTTCAAGCCAACTGCCACCATCAGCACCG 1095
Qy 1156 CTAACGGAAGAGACTTCGGATTAACAAAGCAGCTGCGAACAAGAAAACAAG 1215
Db 1096 CTAACGGAAGAGACTTCGGATTAACAAAGCAGCTGCGAACAAGAAAACAAG 1155
Qy 1216 GCTGTCGAAAGAGAATAATCTTGATGCTGTACACTAGGACTAGCAAGCAG 1275
Db 1156 GCTGTCGAAAGAGAATAATCTTGATGCTGTACACTAGGACTAGCAAGCAG 1215
Qy 1276 CGGTGGC 1282
Db 1216 CGGTGGC 1222

RESULT 4
AR152404
LOCUS AR152404 1541 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 3 from patent US 6232527.
ACCESSION AR152404
VERSION AR152404.1 GI:15118454
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
  1 (bases 1 to 1541)
AUTHORS Mahajan,P.B.
TITLE Maize Rad2/FEN-1 orthologues and uses thereof
JOURNAL Patent: US 6232527-A 3 15-MAY-2001;
FEATURES
  Location/Qualifiers
    source
      1..1541
      /organism="unknown"
BASE COUNT 473 a 308 c 377 g 383 t
ORIGIN

Query Match 57.0%; Score 842; DB 6; Length 1541;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1242; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 33 GCGGTTTCTTGGCCACTTCGGCTCAGCCGCGCGCCGACCAAGCCGCGCAG 92
Db 15 GCGGTTTCTTGGCCACTTCGGCTCAGCCGCGCGCCGACCAAGCCGCGCAG 74
Qy 93 CGAGATGGGCATCAAGGTTTGACGAACTGCTGGCGGACAAATGCGCCCAAGCCGATGAA 152
Db 75 CGAGATGGGCATCAAGGTTTGACGAACTGCTGGCGGACAAATGCGCCCAAGCCGATGAA 134
Qy 153 GGAGCAGAAATTCAGAGCTACTTCGGCCCGCAAAATGCGCGTCGAGCCGAGCATGAGCAT 212
Db 135 GGAGCAGAAATTCAGAGCTACTTCGGCCCGCAAAATGCGCGTCGAGCCGAGCATGAGCAT 194
Qy 213 CTACCACTTCTGTAGTAGTGGAGGACAGGCATGGAACTCTCACAAATGAAGCTGG 272
Db 195 ATACCACTTCTGTAGTAGTGGAGGACAGGCATGGAACTCTCACAAATGAAGCTGG 254
Qy 273 TGAAGTCTACTAGTCAATTTGCAAGAAATGTTCAACCGGACAAATGAATTTACTGGAAGCGGG 332
Db 255 TGAAGTCTACTAGTCAATTTGCAAGAAATGTTCAACCGGACAAATGAATTTACTGGAAGCGGG 314
Qy 333 AATCAAGCCAGTTTATGTTTGTATGGCAAGCCTCTCTGATATGAAGAAACAAGAACTTGC 392
```

```
Db 315 AATCAAGCCAGTTTATGTTTGTATGCAAGCCTCTCTGATATGAAGAAACAAGACTTGC 374
Qy 393 TAAAGATACTCAAAAGACAGATGATGCAACCAAGATCTGACTGAGGCAGTAGAGTAGG 452
Db 375 TAAAGATACTCAAAAGACAGATGATGCAACCAAGATCTGACTGAGGCAGTAGAGTAGG 434
Qy 453 AGATAAAGATCGGATTTGAAAAATTTAGCAAGAGGACTGTAAAGGTCACAAAGCAACACAA 512
Db 435 AGATAAAGATCGGATTTGAAAAATTTAGCAAGAGGACTGTAAAGGTCACAAAGCAACACAA 494
Qy 513 CCAAGATTTGTAACGACTATTAAGACTTTATGGGGTTCCTGTTGTAGAGGCACCTTCTGA 572
Db 495 CCAAGATTTGTAACGCGCTATTAAGACTTTATGGGGTTCCTGTTGTAGAGGCACCTTCTGA 554
Qy 573 AGCAGAAGCAAGATGTGCAGCCCTTTGCATAAACGATAAAGGTTTCGCTGTTGCTTCAGA 632
Db 555 AGCAGAAGCAAGATGTGCAGCCCTTTGCATAAACGATAAAGGTTTCGCTGTTGCTTCAGA 614
Qy 633 AGATAAGGACTCCCTTACTTTTGGGGCTCCACGGTTTCCTTCGTCATTTAATGAATCCAG 692
Db 615 AGATAAGGACTCCCTTACTTTTGGGGCTCCACGGTTTCCTTCGTCATTTAATGAATCCAG 674
Qy 693 TTCCAAGAAAAATACCTGTGATGGAATTTGATTTGCCAAGGTTTGGAGGAGCTTGAAC 752
Db 675 TTCCAAGAAAAATACCTGTGATGGAATTTGATTTGCCAAGGTTTGGAGGAGCTTGAAC 734
Qy 753 CACCATGACGAGTTCATTTGATTTGTCATCTGTCGATGTGACTATTGTATAGGAT 812
Db 735 CACCATGACGAGTTCATTTGATTTGTCATCTGTCGATGTGACTATTGTATAGGAT 794
Qy 813 CAAAGGTATCGGGGGGCAACAGCTCTGAACTTATTCGTCACACATGGGTCCATAGAAG 872
Db 795 CAAAGGTATCGGGGGGCAACAGCTCTGAACTTATTCGTCACACATGGGTCCATAGAAG 854
Qy 873 CATCTTGGAGAACTTTAATAAGACAGATATCAAAATCTTGAGGACTGSCCTTACCAGA 932
Db 855 CATCTTGGAGAACTTTAATAAGACAGATATCAAAATCTTGAGGACTGSCCTTACCAGA 914
Qy 933 AGCTCGAGCTTGTTCAGAGGAGCCTTAATGTGCTCTTTGGATATTTCCCTGAGCTAAAATGGAC 992
Db 915 AGCTCGAGCTTGTTCAGAGGAGCCTTAATGTGCTCTTTGATATTTCCCTGAGCTAAAATGGAC 974
Qy 993 TGCACCTGATGAGGAGGCTCTCATAGTTTCTGTTGTAAGATTAAGTTTCAATGAAGA 1052
Db 975 TGCACCTGATGAGGAGGCTCTCATAGTTTCTGTTGTAAGATTAAGTTTCAATGAAGA 1034
Qy 1053 TCGGTTGCAAGAGCCATAGAGAAGATCAAAATCTGCCAAGAAATAATTCGCAAGGAAG 1112
Db 1035 TCGGTTGCAAGAGCCATAGAGAAGATCAAAATCTGCCAAGAAATAATTCGCAAGGAAG 1094
Qy 1113 ACTCGAGTCTTTTTCAGCCCACTGCCACATCAGCACCGCTAAACGGAAGAGAGAC 1172
Db 1095 ACTCGAGTCTTTTTCAGCCCACTGCCACATCAGCACCGCTAAACGGAAGAGAGAC 1154
Qy 1173 TTCGGATAAACAAGCAAGCAGCTGCGAACAAGAAACAAGGCTGGTGGAAAGAA 1232
Db 1155 TTCGGATAAACAAGCAAGCAGCTGCGAACAAGAAACAAGGCTGGTGGAAAGAA 1214
Qy 1233 ATAACTTGGATGCTTGTATGTACAACTACGAACTACGAAAGCAGCGGTGGC 1282
Db 1215 ATAACTTGGATGCTTGTATGTACAACTACGAAAGCAGCGGTGGC 1264

RESULT 5
AB021666
LOCUS AB021666 1354 bp mRNA linear PLN 25-MAR-2000
DEFINITION Oryza sativa OsFEN-1 mRNA for FEN-1, complete cds.
ACCESSION AB021666
VERSION AB021666.2 GI:4587224
KEYWORDS FEN-1; endonuclease.
SOURCE Oryza sativa cDNA to mRNA.
ORGANISM Oryza sativa
```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (sites)
Kimura, S., Ueda, T., Hatanaka, M., Takenouchi, M., Hashimoto, J. and Sakaguchi, K.
Plant homologue of flap endonuclease-1: molecular cloning, characterization, and evidence of expression in meristematic tissues
Plant Mol. Biol. 42 (3), 415-427 (2000)

JOURNAL
MEDLINE
20256470

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 1354)
Kimura, S., Hashimoto, J. and Sakaguchi, K.
Direct Submission
Submitted (21-DEC-1998) Seisuke Kimura, Science University of Tokyo, Dept. of Applied Biological Sciences, 2641 Yamazaki, Noda, Chiba 278-8510, Japan (E-mail: j36498703@ed.noda.sut.ac.jp, Tel:81-471-24-1501(ex.3419), Fax:81-471-23-9767)
On Apr 17, 1999 this sequence version replaced gi:4062866.

COMMENT
FEATURES
Location/Qualifiers
source
1..1354
/organism="Oryza sativa"
/db_xref="taxon:4530"
67..1209
/gene="osFEN-1"
67..1209
/gene="osFEN-1"
/note="endonuclease"
/codon_start=1
/product="FEN-1"
/protein_id="BAA36171.1"
/db_xref="GI:4587225"
/translation="MGKGLTKLLADNAPKMEKEFSEYGGRIAYDASMSIXOFLI
VVGRTGMEITLNEAGEVTSIHQGFNFTIRLLGAIKPYVFDGPPDLKKVGEIAKRY
SKRDATKELTEAVEEGDKRAIEKFSRTVYKQHNKECKRLRLRMGPVVEAPCEA
EAEACALINDMVAVEDSDMSLTFGAPFLRLMDPSKKIPVNEFEVAKVEELE
LMDQFIDLCISGDCYDSIKGIGQTALKLRQHSIESILEINIKRQVIPEDWP
YQEARLKEFNPVLDIPELKNWPADEGLVEFLVKEGFGNDRVTKALEIKTKAKNK
SSQGRLESFFKPVTSVPLRKDKTSEKPTKAVANKKTKGAGGKKK"
1354
polyA_site
/note="17 A nucleotides"

BASE COUNT 415 a 286 c 331 g 322 t

LOCUS
AC104713
DEFINITION
Oryza sativa chromosome 5 clone OJ1362G11, linear HTG 20-DEC-2001
PROGRESS ***, 3 ordered pieces.
AC104713
AC104713.1 GI:17940800
KEYWORDS
HTG; HTGS_PHASE2.
SOURCE
Oryza sativa
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 115907)
Chow, T.-Y., Hsing, Y.-I. C., Chen, C.-S., Chen, H.-H., Wu, H.-P., Liu, S.-M., Chao, I.-T., Chang, S.-J., Chen, T.-R., Chen, Y.-L., Chow, M.-H. J., Hong, Y.-C., Hsiung, J.-N., Hsu, C.-H., Huang, J.-J., Kau, P.-I., Lee, M.-C., Leu, H.-L., Lin, S.-J., Wu, L.-F. and Shaw, J.-F.
Oryza sativa BAC OJ1362G11 genomic sequence
Unpublished

Query Match 3.9%; Score 58; DB 8; Length 1354;
Best Local Similarity 100.0%; Pred. No. 3.8e-21;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 TGGCGGACAATGCCCAAGCGGATGAGGAGCAGAGCTTCGAGAGCTACTTCGGCG 182
|||||
Db 95 TGGCGGACAATGCCCAAGCGGATGAGGAGCAGAGCTTCGAGAGCTACTTCGGCG 152
|||||

RESULT 6
AC104713/c
LOCUS
DEFINITION
Oryza sativa chromosome 5 clone OJ1362G11, linear HTG 20-DEC-2001
PROGRESS ***, 3 ordered pieces.
AC104713
AC104713.1 GI:17940800
KEYWORDS
HTG; HTGS_PHASE2.
SOURCE
Oryza sativa
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 115907)
Chow, T.-Y., Hsing, Y.-I. C., Chen, C.-S., Chen, H.-H., Wu, H.-P., Liu, S.-M., Chao, I.-T., Chang, S.-J., Chen, T.-R., Chen, Y.-L., Chow, M.-H. J., Hong, Y.-C., Hsiung, J.-N., Hsu, C.-H., Huang, J.-J., Kau, P.-I., Lee, M.-C., Leu, H.-L., Lin, S.-J., Wu, L.-F. and Shaw, J.-F.
Oryza sativa BAC OJ1362G11 genomic sequence
Unpublished

2 (bases 1 to 115907)
Chow, T.-Y. and Hsing, Y.-I. C.
Direct Submission
Submitted (20-DEC-2001) Institute of Botany, Academia Sinica, 128, Section 2, Academia Road, Nankang, Taipei 11529, Taiwan
The nucleotide sequence of this BAC clone was generated by combining Monsanto and ASPG-Taiwan sequencing data.
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 3719: contig of 3719 bp in length
* 3720 3739: gap of unknown length
* 3740 110847: contig of 107108 bp in length
* 110848 110867: gap of unknown length
* 110868 115907: contig of 5040 bp in length.

FEATURES
Location/Qualifiers
source
1..115907
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="5"
/clone="OJ1362G11"
BASE COUNT 32362 a 24964 c 25806 g 32735 t 40 others
ORIGIN

Query Match 3.9%; Score 58; DB 2; Length 115907;
Best Local Similarity 100.0%; Pred. No. 4.5e-21;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 TGGCGGACAATGCCCAAGCGGATGAGGAGCAGAGCTTCGAGAGCTACTTCGGCG 182
|||||
Db 8213 TGGCGGACAATGCCCAAGCGGATGAGGAGCAGAGCTTCGAGAGCTACTTCGGCG 8156
|||||

RESULT 7
AL590807
LOCUS
DEFINITION
Human DNA sequence from clone Rp11-325L2 on chromosome 13, complete sequence.
AL590807
AL590807.13 GI:18642375
KEYWORDS
HTG.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 178914)
Tromans, A.
Direct Submission
Submitted (07-FEB-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Feb 8, 2002 this sequence version replaced gi:18614047.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em., EMBL, SW., SWISSPROT, Tr., TREMBL, Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr13>
 RP11-325L2 is from the library RP11-11.2 constructed by the group of Pleter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pPACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-325L2 it may be shorter because we sequence overlapping sections only once, except for a short overlap.
 The true left end of clone RP11-325L2 is at 1 in this sequence. The true left end of clone RP11-521J24 is at 176915 in this sequence. The true right end of clone RP11-454P4 is at 87814 in this sequence.

```

FEATURES             Location/Qualifiers
     source           1..178914
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /clone="RP11-325L2"
                     /clone_lib="RPCI-11.2"
     misc_feature     55682..55741
                     /note="Sequence from overlapping clone RP11-454P4
                     (AL157696). Assembly confirmed by
                     restriction digest."
     misc_feature     134535
                     /note="Tandem repeat. Forced join. Gap size estimated to
                     be approximately 217bp by restriction digest data."
     misc_feature     134536..134656
                     /note="Single clone region. Assembly consistent with
                     restriction digest data."
     BASE COUNT      55984 a 30899 c 31660 g 60371 t
     ORIGIN

```

```

Query Match          2.2%; Score 32; DB 9; Length 178914;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1447 ATCCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
|||||
Db 72351 ATGCCACTAAATAAAAAAAAAAAAAAAAAAAAAA 72382

```

```

RESULT 8
AX147252
LOCUS              851 bp DNA linear PAT 08-JUN-2001
DEFINITION         Sequence 39 from Patent WO0136584.
ACCESSION          AX147252
VERSION            AX147252.1 GI:14346422
KEYWORDS
SOURCE             human.
ORGANISM           Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 851)
Cunningham,M.J., Zweiger,C.B., Kaser,M.R., Panzer,S.R.,
Seilhammer,J.J., Yue,H., Baughn,M.R., Azimzai,Y. and Lal,P.
Mammalian toxicological response markers
Patent: WO 0136684-A 39 25-MAY-2001;
Incyte Genomics, Inc. (US)

```

```

FEATURES             Location/Qualifiers
     source           1..851
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No.: 3123954"
     BASE COUNT      176 a 340 c 214 g 119 t
     ORIGIN

```

```

Query Match          2.0%; Score 30; DB 6; Length 851;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1449 GCCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
|||||
Db 806 GCCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 835

```

```

RESULT 9
AK027130
LOCUS              1487 bp mRNA linear PRI 29-SEP-2000
DEFINITION         Homo sapiens cDNA: FLJ23477 fis, clone HS115732.
ACCESSION          AK027130
VERSION            AK027130.1 GI:10440180
KEYWORDS           oligo capping; fis (full insert sequence).
SOURCE             Homo sapiens human small intestine cDNA to mRNA, clone_lib:HSI
clone:HS115732.

```

```

ORGANISM           Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished (2000)

```

```

REFERENCE          2 (bases 1 to 1487)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdnaledms.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Departent of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

```

```

FEATURES             Location/Qualifiers
     source           1..1487
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HS115732"
                     /tissue_type="human small intestine"
                     /clone_lib="HSI"
                     /note="cloning vector pME18SFL3"
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB15669.1"
                     /db_xref="GI:10440181"
                     /translation="MLSCLEPAAWDCGAEGRDREEMVLGWRKDRGTTPTPLPASPS
                     ACTPLWAGHVLGDAGSEPLSPHPCPNRSCWRSKSCGLAGWRCGSEAAPLTPPE
                     VGLRQDPGVGRNPSILPAPQPVNRLTFCGPAPASDRP"
     CDS              278 a 427 c 470 g 312 t

```

```

BASE COUNT          278 a 427 c 470 g 312 t
ORIGIN
Query Match          2.0%; Score 30; DB 9; Length 1487;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1449 GCCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
|||||
Db 1456 GCCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1485

```

```

RESULT 10
BC020218

```



```

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 110811)
Nakamura,Y.
Direct Submission
Submitted (13-DEC-2001) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:ynakam@kazusa.or.jp,
URL:http://www.kazusa.or.jp, Tel:81-438-52-3935,
Fax:81-438-52-3934)

FEATURES
source
location/Qualifiers
1..110811
/organism="Lotus japonicus"
/db_xref="taxon:34305"
/chromosome="5"
/clone="LJT05P01"
/clone_lib="LJT library"
/note="TAC clone:TM0048"

BASE COUNT
35030 a 19786 c 20743 g 35252 t

ORIGIN
Query Match 2.0%; Score 30; DB 8; Length 110811;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 TTGCAGCTACTTCGGCGCAAAATCGCC 192
|||||
Db 78321 TTGCAGCTACTTCGGCGCAAAATCGCC 78292

RESULT 13
AC103556
LOCUS
DEFINITION
Rattus norvegicus chromosome Chr18 clone CH230-24C4, *** SEQUENCING
IN PROGRESS ***, 74 unordered pieces.
AC103556
AC103556.2 GI:17975674
HTG: HTGS_PHASE1.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 133139)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Arc,J.R., Banks,T., Barbaria,J.,
Benton,J., Binage,M., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinu,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flag,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,
Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,
Ogub,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,

```

```

Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Umani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 133139)
Worley,K.C.
Direct Submission
Submitted (28-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Raylor Plaza, Houston, TX 77030, USA
On Dec 21, 2001 this sequence version replaced gi:17136085.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GSKO
Center clone name: CH230-24C4
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findPhrapList
Consensus quality: 106257 bases at least Q40
Consensus quality: 115409 bases at least Q30
Consensus quality: 122085 bases at least Q20
Estimated insert size: 81889; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 1x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 74 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1
* 3373: contig of 3373 bp in length
* 3374: gap of unknown length
* 3474: contig of 3499 bp in length
* 6973: gap of unknown length
* 7073: contig of 4344 bp in length
* 11417: gap of unknown length
* 11517: contig of 1952 bp in length
* 13469: gap of unknown length
* 13569: contig of 2992 bp in length
* 16561: gap of unknown length
* 16661: contig of 2215 bp in length
* 18876: gap of unknown length
* 18975: contig of 1642 bp in length
* 20617: gap of unknown length
* 20718: contig of 2152 bp in length
* 22870: gap of unknown length
* 22969: contig of 2787 bp in length
* 25757: gap of unknown length
* 25856: contig of 2761 bp in length
* 28618: gap of unknown length
* 30131: contig of 1414 bp in length
* 30232: gap of unknown length
* 32843: contig of 2612 bp in length
* 32843: gap of unknown length
* 34994: contig of 1951 bp in length
* 34995: gap of unknown length
* 37183: contig of 2189 bp in length
* 37184: gap of unknown length

```



```

130097. .130287.133310. .133431.136357. .136475.
139133. .139335.141728. .141801.144155. .144307.
145204. >145359)
/product="axonemal dynein heavy chain 8 Dnahe8"
join(3069. .3256.9629. .9709.10121. .10237.12267. .12386.
16057. .16192.19820. .19932.20981. .21087.26266. .26338.
27737. .27923.30621. .30817.32357. .32483.34263. .34498.
36464. .36568.38235. .38348.38996. .39058.41443. .41543.
43704. .43841.44900. .45140.47076. .47275.48975. .49128.
51158. .51321.52830. .52932.54211. .54315.55129. .55290.
57068. .57292.62035. .62207.63299. .63477.65616. .65803.
67175. .67261.67356. .67463.68704. .68838.70358. .70719.
77590. .77681.78140. .78240.78882. .74022.74543. .74780.
77539. .77306.77319. .73424.73862. .79033.81058. .81271.
82746. .82855.84288. .84451.85609. .85801.87045. .87167.
88668. .88798.91627. .91780.94035. .94199.95873. .96099.
98024. .98142.99756. .99859.101154. .101330.105215. .105382.
105931. .106146.108588. .108725.110279. .110406.
111223. .111394.113361. .113516.117047. .117202.
120848. .121030.126557. .126662.129154. .129406.
130097. .130287.133310. .133431.136357. .136475.
139133. .139335.141728. .141801.144155. .144307.
145204. >145359)
/notes="ATPase: putative motor unit of axonemal dynein(s)
in mouse sperm flagellum; has been implicated in both the
regulation of motility and flagellar assembly in Mus
musculus sperm; candidate for t complex
distorter/sterility gene 2 and hybrid sterility 6 gene;
synonym: DLP8; similar to the products encoded by GenBank
Accession Numbers AF356519, AF356520, AF356521, AF356522,
AF356523, AL0391415, AL035555, AC022402, AL034345,
AL353700, AL035690, AJ132090, AJ132091"
/codon_start=3
/product="axonemal dynein heavy chain 8 Dnahe8"
/db_xref="GI:13310482"
/translation="IQATLVRHPETGKLVNFDPKILEVVRFTKCMIKMKLDVPEQA
KNLLKLEKADKLYLQGLDYDDICQEPVSVNLMTPRMKKVSVESLROGLVLT
WSSLMSEFKEKADYVLDNMFNOLKLVKNCEDHIDTVLKEIAKTLLISLSGGTIVT
EDVLGNETEYTKCADLNHKSRRHVEAEVLELIFEQIYEVKTYCKAASKVKEQKRK
VFMGSEAEYEGLDFTSTMEVTDNDEDEKFEKVEYAFVTSQHLDSLOKATRLSL
DMKRRIFQGGRRSEIVSRKTEVHLAIPNVYVPSLDDIOQAINRMLQTLFV
SRGVHMGQOVRQIKSFONNSQDOPASGPKLKEPSEETIPARKLNFVGV
AEHDKISLVLSSVNSLRKAATEALQDFQKYKTLWIEDRHVKVKEFLANPSLTE
IRSELTHLLEQIEKELPIVVGSLHTEPMKLSLSEAKAMKMLCLPINEEYIL
KMSDMITFINELKLSRPIRDLDDVFAMEALSCIRDETQMDTLGPIBEAYGIL
NRFVKYTKESGVDTRYSFNKLQKSVQOGLVQVQPKFSNLLSVKVFCEDY
INFTEAYTEGPMVNPPOEASNRQIFQANEDDLWRKFVTVSSGEQLFGPVTDYE
VLHTRKLENLQKLYGLDYTVMGSTSGYVEILGVDVIEKINAELOEQFNRCKLPR
ALKDQWAFDLKRRIDFSESCPLEMNTNKKORWORISETCTPFDVSDTFCL
PNIWEAPLKNKDDIEDICISAKEDIEAKLQVIENTYQNLSPAEEKGELLK
GTSEGIITLMEDSLVGLSLLNRYNTPKKNQWVFKLSTSDIIEWLVNQLW
VYLEAVFVGSDIAQLPOEAKRFQNDKWKIMQRAHENPNVISCVDGEMGQLLP
HLHLEQVCKRSLGYLEKRLFPREFVSDVLEILQASDSHTIQPHLPAYSDN
INFTVHAKQYDRMTAVISREKIMLDPVMAKGPVEIWLDDLVKQVSSLHNIIRS
AFYQISDGSFLLPLFHPAQVGLLQIHWORIDHSEALNNAKDKTMQITNOKFL
DILNTLSOHTDITKEDRVKFTLTIHVHORDIFDLVAMHISKVDFEMLKOSRF
YEKEDLQTVMSITDVFYONBELGCTDLVLTPTDRCYITLQALGMNMGAGPAG
PAGTGTETKMGRCGLGKIVFVNCSDQMDRGLGRIFNGLAQSSGSGDFEFNRKE
LPVLSAAQQLIYVLTARKKKQFIFSDGDCVDLNPFEFGIFITMNPAGQROEPEN
LKIFRTVAMVPDQRIIMRVKLASCGLFENVLQAKFYVLYKLCQEQLTKQHVYDGL
LNLIVLTLGSKQKARPEDSELSTVMGRIDMNI.SKLVDEDEPFLSLINDLFPFLG
QLDSTVAELQASVDNQVNEGLINHPNKLVLQVETSLVRHGLMTLGPVSGSKGT
VITILMSLTCFGRPHRMNPKAITAPMGRGLDTATNDTGDFTGLSTLRKTKLAK
KGENIFLIDGPDAIENLNSVLDNKTLLTANGDIRPMACTCKLLEVENILENAS
PATVSRMGVYISSALSWRPIQLAWLKRKQOASVFLISDKKVFEDAYTYMKLSLN
PMQOLLECINLLEGLTPSEKGGVSSGDHLKUFVFGMLMSGLALFLDSRE
KLEVFLRHGSKLNLPEIPKSGOQTYEFYVTDGWEHWNKRRIQFYFPTDIPSEYS
SILVFNVDINTNLNEDITIAOKHALLTGEOTAKTVMYKAYLKRYDEVLQSKLS
FSSATPEMFOFTESVDKRMGSTVGPGRKMTVFIDINMPVINWEGDQITNEIV
ROMMEMEGM.SLDGQFDTTIVDQLIAAMHPGGRNDIPQRLKQFQVFNCTFNIF
TSIDKIFIGICGQYDFCRFRPEICDMGNLVSRSVLWQTKVWLPTPKSFHYIF
NLRLSLRWQMLTVKAECCSIPILLSLFKECNRVIAIDRFTTPDDEQWNSQLIRA

```

```

VERNISPEVAANILPEPYFVDFLDRMPETGDEPTMTPEVKIYELVPSFEFLSEKL
QFYOROFNETIRGTSLDLVFFXDMTHLVKISRIITSCGNALLVGVGSGSGKSLK
ASLQAGYQIQTILTRSYNVSNLIEDKLVKAVAGKGTIFTFDNEIKDEAFLEY
LNNLSGSEISNLFARDEMDLTOGLISYMKBELPHPTEDNLYEITRSRKNLHV
VLCPSPGKFRASRLKFCGLISGCTMDFSWRKPKEALIAVASYELLYDYNVCSLETX
RHVETVGFHDWVSESCNYFQYRRRAHVTPKSYLSFINGYKISYITDKVKYINEQA
ERNMIGLIDKIMEASESVAKLSODLAVKEKELASIKADEVLAETVSAQASAKVNE
VQVQKQAKIIVDEIDSEKVKABTKLEAAKPALEAEALNTIKPNDIATVKLAKPP
HLIMRIMDCVLLIFOKKIDPVTMDPKCKPSGWSGSKMSATGFLFSLQOFPKDTI
NEETVELLOPYFNMDYTTFESAKVCGVAGLLSWTLMVIFYGINREVLPKANLAK
OEGRLVANVFLKQAALLDEKQALDKVQAFDAAMEKMDANDAMCRKMOAAS
TLIDLSGSEKVRTQSKSEKQINRLVGDVLLCTGFLSYLQFPFNOIFERNYLLKDOWE
LEKARKIPETENLNIAMLDVPTIGEMGLOGLPDDUSIONGLIIVATKRIPLDIL
POTQGTWTKSKREKENDLQVTSINHRIFRTHLEDSLSUGRPLLIEDIREDLPALDN
LEKNFTKSGTAFKVKDEKEDIMDTFKLYITKLPNPAFTPEINAKTSVIDFTVMK
GLNQNLRLVILFKQELSERVLEKEDFTFKRKMKELEDNLLYKLSATGSLVDDE
SLIGLRITKOTAAEYSEKLVHAAETIEIKINTAQEFEPAPATGSLYFLITENSMVN
IMYOTSLAQTKLFDOSMARSEKSPUPOKRITNILEYTYEVETYSVGLVENHKLFL
VILMTLKIDLORGTVKHKSFOALIKGGAOLDKACPPKFRWILOMTNLNVLKSLP
OFAEIMNQISRNKSGKWNFDKDAPEEIIPOYNDSLDTCRKLILIRSWCDPRTVFO
ARKEIMNQISRNKSGKWNFDKDAPEEIIPOYNDSLDTCRKLILIRSWCDPRTVFO
ARKYIADSLSEKYTEPVLNLEKTWEESDTHPLICFLSMGSDPTQIDALAKKULE
CRTISMGQGVHARKLIQLSMOQGGVLLQNLCHLGEFMEELMEMAMVETTESDFR
WITTEPHDRFPITLLQTSIKFTNEPPQGVRAGLKRTFAGINQDILLDISNLPMMKPMIL
YTVAFLHSTVO"
BASE COUNT 40485 a 33581 c 34177 g 41921 t
ORIGIN

```

Query Match 2.0%; Score 30; DB 10; Length 150164;
 Best Local Similarity 100.0%; Pred. No. 2.6e-05;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1449 GCACGTAAAAA...AAAAAAAAAAAA 1478

Db 126747 GCACGTAAAAA...AAAAAAAAAAAA 126718

RESULT 15
 AC024152/c
 LOCUS
 DEFINITION Homo sapiens chromosome 3 clone RP11-24E1, WORKING DRAFT SEQUENCE,
 15 unordered pieces.
 AC024152
 AC024152.11 GI:11079270
 VERSION
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 162201)
 AUTHORS Muzny D.M., Adams C., Adio-Oduola B., Ali-osman, F.R., Allen C.,
 Albrooks S.L., Amaratunge, H.C., Are, J.R., Banks, T., Barbara, J.,
 Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J.,
 Bowie, S., Brivava, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
 Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
 Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R.,
 Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,
 Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
 Davy-carroll, L., Dederich, D.A., Delaney, K.R., Delgado O.,
 Deun, A.L., Ding, Y., Dinh, H.H., Douchwaite, K.J., Drapet, H.,
 Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,
 Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,
 Foster, P., Frantz, G., Gabisi, A., Gao, J., Garcia, A., Garner, T.,
 Garza, N., Gill, R., Gorell, J.H., Guevara, W., Gunaratne, P., Hale, S.,
 Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,
 Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C.,
 Hollins, B., Homsai, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
 Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
 Joudas, S., Karlovic, E., Kelly, S., Khan, U., King, L., Korvah, J.,
 Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, J.C.,
 Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
 Louis, L., Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
 Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
 Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M.,


```

/standard_name="D12S2444"
/db_xref="GDB:978593S"
complement(17717..18009)
/rpt_family="AluJo"
complement(19048..19436)
/rpt_family="Alu_rich"
19140..19790
/function="Low coverage"
19646..19802
/rpt_family="AluSg/x"
complement(19805..19843)

Query Match          2.0%; Score 30; DB 9; Length 191134;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1449 GCCAGTAAAAA 1478
|||||
Db 76531 GCCAGTAAAAA 76560

RESULT 17
LOCUS          AC020708      208360 bp      DNA      linear      PRI 30-MAY-2001
DEFINITION     Homo sapiens clone RP11-425B17, complete sequence.
ACCESSION      AC020708
VERSION        AC020708
KEYWORDS       HTG.
SOURCE         human.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 208360)
AUTHORS        Waterston,R.H.
TITLE          The sequence of Homo sapiens clone
JOURNAL         Unpublished
REFERENCE      2 (bases 1 to 208360)
AUTHORS        Waterston,R.H.
TITLE          Direct Submission
JOURNAL         Submitted (08-JAN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE      3 (bases 1 to 208360)
AUTHORS        Waterston,R.H.
TITLE          Direct Submission
JOURNAL         Submitted (30-MAY-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT        On May 30, 2001 this sequence version replaced gi:14018139.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0425B17

FEATURES
source          1..208360
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="RP11-425B17"

BASE COUNT      62719 a 38793 c 40657 g 66191 t
ORIGIN

Query Match          2.0%; Score 30; DB 9; Length 208360;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1449 GCCAGTAAAAA 1478
|||||
Db 105616 GCCAGTAAAAA 105645

RESULT 18
AC096327/c
LOCUS          AC096327      233048 bp      DNA      linear      HTG 20-DEC-2001
DEFINITION     Rattus norvegicus chromosome rfl clone CH230-72M16, WORKING DRAFT
SEQUENCE, 20 unordered pieces.
AC096327
AC096327.4 GI:17944026
HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 233048)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.I., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,K., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseyed,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meadot,M.,
Mei,G., Metsker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,A., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,M., Nguyen,N., Nguyen,N., Nickerson,E., Nwokoko,S.,
Ogih,M., Okuwona,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N.,
Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 233048)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:15799432.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GEXY
Center clone name: CH230-72M16
----- Summary Statistics -----
Assembly program: Phrap; version 0.990329First call to
findPhrapList

```

```
Consensus quality: 225714 bases at least Q40
Consensus quality: 228399 bases at least Q30
Consensus quality: 230469 bases at least Q20
Estimated insert size: 224958; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; agarose-gel estimation
Quality coverage: 6.6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  * consists of 20 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
*
* 1 52974: contig of 52974 bp in length
* 52975 53074: gap of unknown length
* 53075 79799: contig of 26725 bp in length
* 79800 79800: gap of unknown length
* 79900 101381: contig of 21482 bp in length
* 101382 101481: gap of unknown length
* 101482 122794: contig of 21313 bp in length
* 122795 122894: gap of unknown length
* 122895 123670: contig of 15776 bp in length
* 123671 138671: gap of unknown length
* 138672 152605: contig of 13835 bp in length
* 152606 152705: gap of unknown length
* 152706 162732: contig of 10027 bp in length
* 162733 162832: gap of unknown length
* 162833 173380: contig of 10548 bp in length
* 173381 173480: gap of unknown length
* 173481 184068: contig of 10588 bp in length
* 184069 184168: gap of unknown length
* 184169 193398: contig of 9230 bp in length
* 193400 193499: gap of unknown length
* 193500 201256: contig of 7758 bp in length
* 201257 201356: gap of unknown length
* 201357 209184: contig of 7828 bp in length
* 209185 209284: gap of unknown length
* 209285 214977: contig of 5693 bp in length
* 214978 215077: gap of unknown length
* 215078 220792: contig of 5715 bp in length
* 220793 220892: gap of unknown length
* 220893 224492: contig of 3600 bp in length
* 224493 224592: gap of unknown length
* 224593 226693: contig of 2101 bp in length
* 226694 226793: gap of unknown length
* 226794 228673: contig of 1880 bp in length
* 228674 228773: gap of unknown length
* 228774 230043: contig of 1270 bp in length
* 230044 230143: gap of unknown length
* 230144 231574: contig of 1431 bp in length
* 231575 231674: gap of unknown length
* 231675 233048: contig of 1374 bp in length.
*
* Location/Qualifiers
*     1..233048
*         /organism="Rattus norvegicus"
*         /db_xref="taxon:10116"
*         /chromosome="Rfl"
*         /clone="CH230-72M16"
*
* BASE COUNT  67063 a 48857 c 48194 g 66783 t 2151 others
* ORIGIN
*
* Query Match      2.0%; Score 30; DB 2; Length 233048;
* Best Local Similarity 100.0%; Pred. No. 2.7e+05;
* Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
*
* QY 1449 GCCAGTAAAAAAGAAAAAAGAAAAAAGAAAAA 1478
*      |||||||
* Db 144671 GCCAGTAAAAAAGAAAAAAGAAAAAAGAAAAA 144642
```

```
RESULT 19
AC090867/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AC090867.2 GI:14140274
HTG: HTGS_PHASE1; HTGS_DRAFT.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Smith, D.R.
Genome Therapeutics Corporation Sequencing Center: Mouse Genome
Sequence Data
Unpublished
2 (bases 1 to 244849)
Smith, D.R.
Direct Submission
Submitted (16-MAR-2001) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
On May 17, 2001 this sequence version replaced gi:13357224.
-----
Center: Genome Therapeutics Corporation
Center code: GTC
Web site: http://www.genomecorp.com/
Contact: gtc-seqcenter@genomecorp.com
----- Project Information
Center project name: mg045
----- Summary Statistics
Sequencing vector: N/A
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 990315
Consensus quality: 205883 bases at least Q40
Consensus quality: 212730 bases at least Q30
Consensus quality: 217320 bases at least Q20
Insert size: 239048; sum-of-contigs
Quality coverage: 6.3x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 60 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1087: contig of 1087 bp in length
* 1088 1187: gap of unknown length
* 1188 2310: contig of 1123 bp in length
* 2311 2410: gap of unknown length
* 2411 3565: contig of 1155 bp in length
* 3566 3665: gap of unknown length
* 3666 4669: contig of 1004 bp in length
* 4670 4769: gap of unknown length
* 4770 5801: contig of 1032 bp in length
* 5802 5901: gap of unknown length
* 5902 6980: contig of 1079 bp in length
* 6981 7080: gap of unknown length
* 7081 8101: contig of 1021 bp in length
* 8102 8201: gap of unknown length
* 8202 9263: contig of 1062 bp in length
* 9264 9363: gap of unknown length
* 9364 10418: contig of 1055 bp in length
* 10419 10518: gap of unknown length
* 10519 11564: contig of 1046 bp in length
* 11565 11864: gap of unknown length
* 11865 12695: contig of 1031 bp in length
* 12696 12795: gap of unknown length
* 12796 13837: contig of 1042 bp in length
* 13838 13937: gap of unknown length
```

```
* 13938 15262: contig of 1325 bp in length
* 15263 15362: gap of unknown length
* 15363 16470: contig of 1108 bp in length
* 16471 16570: gap of unknown length
* 16571 17578: contig of 1008 bp in length
* 17579 17678: gap of unknown length
* 17679 18825: contig of 1147 bp in length
* 18826 18925: gap of unknown length
* 18926 20392: contig of 1467 bp in length
* 20393 20492: gap of unknown length
* 20493 22063: contig of 1571 bp in length
* 22064 22163: gap of unknown length
* 22164 23706: contig of 1543 bp in length
* 23707 23806: gap of unknown length
* 23807 24854: contig of 1048 bp in length
* 24855 24954: gap of unknown length
* 24955 26081: contig of 1127 bp in length
* 26082 27426: gap of unknown length
* 27427 27526: gap of unknown length
* 27527 28552: contig of 1326 bp in length
* 28553 28952: gap of unknown length
* 28953 30052: contig of 1100 bp in length
* 30053 30152: gap of unknown length
* 30153 31941: contig of 1789 bp in length
* 31942 32041: gap of unknown length
* 32042 33216: contig of 1175 bp in length
* 33217 33316: gap of unknown length
* 33317 34737: contig of 1421 bp in length
* 34738 34837: gap of unknown length
* 34838 36363: contig of 1526 bp in length
* 36364 36463: gap of unknown length
* 36464 38108: contig of 1643 bp in length
* 38109 38208: gap of unknown length
* 38209 39511: contig of 1305 bp in length
* 39512 39611: gap of unknown length
* 40971 41071: contig of 1360 bp in length
* 40972 42429: contig of 1358 bp in length
* 42430 42529: gap of unknown length
* 42530 43545: contig of 1016 bp in length
* 43546 43645: gap of unknown length
* 43646 45168: contig of 1524 bp in length
* 45169 45268: gap of unknown length
* 45270 46563: contig of 1296 bp in length
* 46564 47874: contig of 1209 bp in length
* 47875 49206: contig of 1232 bp in length
* 49207 49306: gap of unknown length
* 49307 50765: contig of 1459 bp in length
* 50766 50865: gap of unknown length
* 50866 52468: gap of unknown length
* 52469 54162: contig of 1694 bp in length
* 54163 54262: gap of unknown length
* 54263 56430: contig of 2168 bp in length
* 56431 56530: gap of unknown length
* 56531 58280: contig of 1750 bp in length
* 58281 59380: gap of unknown length
* 59381 59761: contig of 1381 bp in length
* 59762 59861: gap of unknown length
* 59862 61591: contig of 1730 bp in length
* 61592 63363: contig of 1672 bp in length
* 63364 63463: gap of unknown length
* 63464 65672: contig of 2209 bp in length
* 65673 65772: gap of unknown length
* 65773 67127: contig of 1355 bp in length
* 67128 68621: contig of 1394 bp in length
* 68622 68721: gap of unknown length
* 68722 69736: contig of 1015 bp in length
```

```
* 69737 69836: gap of unknown length
* 69837 71299: contig of 1463 bp in length
* 71300 71399: gap of unknown length
* 71400 73412: contig of 2013 bp in length
* 73413 73512: gap of unknown length
* 73513 75262: contig of 1750 bp in length
* 75263 75362: gap of unknown length
* 75363 78498: contig of 3136 bp in length
* 78499 78598: gap of unknown length
* 78599 79711: contig of 1113 bp in length
* 79712 79811: gap of unknown length
* 79812 82466: contig of 2655 bp in length
* 82467 82566: gap of unknown length
* 82567 84608: contig of 2042 bp in length
* 84609 84708: gap of unknown length
* 84709 107866: contig of 23158 bp in length
* 107867 107967: gap of unknown length
* 107968 150738: contig of 42772 bp in length
* 150739 150838: gap of unknown length
* 150839 195573: contig of 44735 bp in length
* 195574 195674: gap of unknown length
* 195675 244849: contig of 49176 bp in length.
```

FEATURES

Source

```
Location/Qualifiers
1..244849
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RP23-404D8"
/clone_lib="RPCI-23"
1..1087
/misc_feature
/feature="assembly_name:Contig5"
1188..2310
/misc_feature
/feature="assembly_name:Contig9"
2411..3565
/misc_feature
/feature="assembly_name:Contig10"
3666..4669
/misc_feature
/feature="assembly_name:Contig11"
4770..5801
/misc_feature
/feature="assembly_name:Contig13"
5902..6980
/misc_feature
/feature="assembly_name:Contig14"
7081..8101
/misc_feature
/feature="assembly_name:Contig15"
8202..9283
/misc_feature
/feature="assembly_name:Contig16"
9364..10418
/misc_feature
/feature="assembly_name:Contig20"
10519..11564
/misc_feature
/feature="assembly_name:Contig21"
11665..12695
/misc_feature
/feature="assembly_name:Contig32"
12796..13837
/misc_feature
/feature="assembly_name:Contig33"
13938..15262
/misc_feature
/feature="assembly_name:Contig39"
15363..16470
/misc_feature
/feature="assembly_name:Contig41"
```

Query Match

2.0% Score 30; DB 2; Length 244849;

Best Local Similarity 100.0%; Pred. No. 2.7e-05; Matches 0; Indels 0; Gaps 0;

QY 1448 TCCACGTAAAAAATAAAAAAAAAAAAAA 1477
|||||AAAAAAAAAAAAAAAAAAAAAAAAAAAA
Db 194531 TGCACGTAAAAAATAAAAAAAAAAAAA 194502

RESULT 20

BC015361 BC015361 372 bp mRNA linear PRI 04-OCT-2001
LOCUS Homo sapiens, clone IMAGE:3886095, mRNA.
DEFINITION BC015361
ACCESSION BC015361
VERSION BC015361.1 GI:15929880

```

KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens

REFERENCE
AUTHORS     Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
TITLE       1 (bases 1 to 372)
JOURNAL     Strausberg, R.

REMARK
COMMENT     Direct Submission
            Submitted (01-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC) Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
            NIH-MGC Project URL: http://mgc.nci.nih.gov
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue procurement: DCTD/DP/Genetics
            cDNA Library preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA 94305
            Web site: http://www.shgc.stanford.edu
            Contact: (Dickson, Mark) md@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
            R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAK Plate: 22 Row: c Column: 19.

FEATURES
source      Location/Qualifiers
            1..372
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:3886095"
            /tissue_type="Lung, carcinoma, large cell
            undifferentiated."
            /clone_lib="NIH-MGC_69"
            /lab_host="DH10B"
            /note="vector: pcwv-sport6"

BASE COUNT  195 a 44 c 44 g 89 t
ORIGIN
Query Match      2.0%; Score 29; DB 9; Length 372;
Best Local Similarity 100.0%; Pred. No. 7.7e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAAAA...AAAAAAAAAAAAA 1478
|||||
DB 257 CCAGTAAAAA...AAAAAAAAAAAAA 285

RESULT 21
CAR010224      871 bp mRNA linear PLN 07-AUG-1998
LOCUS          Cicer arietinum mRNA for glyceraldehyde 3-phosphate dehydrogenase,
DEFINITION     partial.
ACCESSION      AJ010224
VERSION        AJ010224.1 GI:3413164
KEYWORDS       cytosolic; gapdh gene; glyceraldehyde 3-phosphate dehydrogenase;
SOURCE         chickpea.
ORGANISM       Cicer arietinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae;
Cicer.
            1 (bases 1 to 871)
            Direct Submission
            Submitted (05-AUG-1998) Labrador E., Fisiologia Vegetal, Univ.
            Salamanca, Campus Miguel de Unamuno. Pza. Doctores de la Reina s/n,
            37007, SPAIN
            2 (bases 1 to 871)
            Dopicco, B., Munoz, F.J. and Labrador, E.

KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens

REFERENCE
AUTHORS     Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
TITLE       1 (bases 1 to 372)
JOURNAL     Strausberg, R.

REMARK
COMMENT     Direct Submission
            Submitted (01-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC) Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
            NIH-MGC Project URL: http://mgc.nci.nih.gov
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue procurement: DCTD/DP/Genetics
            cDNA Library preparation: Life Technologies, Inc.
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA 94305
            Web site: http://www.shgc.stanford.edu
            Contact: (Dickson, Mark) md@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
            R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAK Plate: 22 Row: c Column: 19.

FEATURES
source      Location/Qualifiers
            1..372
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:3886095"
            /tissue_type="Lung, carcinoma, large cell
            undifferentiated."
            /clone_lib="NIH-MGC_69"
            /lab_host="DH10B"
            /note="vector: pcwv-sport6"

BASE COUNT  195 a 44 c 44 g 89 t
ORIGIN
Query Match      2.0%; Score 29; DB 9; Length 372;
Best Local Similarity 100.0%; Pred. No. 7.7e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAAAA...AAAAAAAAAAAAA 1478
|||||
DB 257 CCAGTAAAAA...AAAAAAAAAAAAA 285

RESULT 21
CAR010224      871 bp mRNA linear PLN 07-AUG-1998
LOCUS          Cicer arietinum mRNA for glyceraldehyde 3-phosphate dehydrogenase,
DEFINITION     partial.
ACCESSION      AJ010224
VERSION        AJ010224.1 GI:3413164
KEYWORDS       cytosolic; gapdh gene; glyceraldehyde 3-phosphate dehydrogenase;
SOURCE         chickpea.
ORGANISM       Cicer arietinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae;
Cicer.
            1 (bases 1 to 871)
            Direct Submission
            Submitted (05-AUG-1998) Labrador E., Fisiologia Vegetal, Univ.
            Salamanca, Campus Miguel de Unamuno. Pza. Doctores de la Reina s/n,
            37007, SPAIN
            2 (bases 1 to 871)
            Dopicco, B., Munoz, F.J. and Labrador, E.

TITLE       Glyceraldehyde 3-phosphate dehydrogenase mRNA expressed in chickpea
            epicotyls
JOURNAL     Unpublished
FEATURES     source      Location/Qualifiers
            1..871
            /organism="Cicer arietinum"
            /cultivar="Castellana"
            /db_xref="taxon:3827"
            /clone="CanGAPC"
            /tissue_type="etiolated epicotyls"
            /clone_lib="CAN-5"
            /dev_stage="5 day old seedling"
            1..586
            /gene="gapdh"
            <1..>586
            /gene="gapdh"
            /EC_number="1.2.1.12"
            /function="glycolysis"
            /codon_start=3
            /product="glyceraldehyde 3-phosphate dehydrogenase,
            cytosolic"
            /protein_id="CAA09040.1"
            /db_xref="GI:3413165"
            /db_xref="SPTREMBL:O81924"
            /translation="EYKPEFDIVSNASCTTNCIAPLAKVINDRFGIVELMTTWHST
            ATQKTVDGSAKMRGGRASFNIPSTGAARAVGKVLPLNGKLTGMARFPTVDV
            SVVDLTVKLEKATYDEIKAAINESEGLKGLGYTEDDVYSTDFIGDTRSSIFDAK
            AGIALNDRFKVLSWFDNELGYSTRVVDLVHIA"

polyA_signal  817..822
BASE COUNT    233 a 171 c 200 g 267 t
ORIGIN
Query Match      2.0%; Score 29; DB 8; Length 871;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAAAA...AAAAAAAAAAAAA 1478
|||||
DB 840 CCAGTAAAAA...AAAAAAAAAAAAA 868

RESULT 22
AX022319      1018 bp DNA linear PAT 07-SEP-2000
LOCUS          Sequence 39 from Patent EP0948522.
DEFINITION     AX022319
ACCESSION      AX022319
VERSION        AX022319.1 GI:10045966
KEYWORDS       unidentified.
SOURCE         unidentified.
ORGANISM       unidentified.
REFERENCE      1 (bases 1 to 1018)
AUTHORS        Viney, E.M., Starr, R., Metcalf, D., Nicola, N.A., Willson, T.A.,
            Alexander, W.S., Hilton, D.J., Nicholson, S.E., and Richardson, R.T.
            Therapeutic and diagnostic agents capable of modulating cellular
            responsiveness to cytokines
            Patent: EP 0948522-A 39 13-OCT-1999;
            INST MEDICAL W & E HALL (AU)
            Location/Qualifiers
            1..1018
            /organism="unidentified"
            /db_xref="taxon:32644"

BASE COUNT      306 a 198 c 211 g 303 t
ORIGIN
Query Match      2.0%; Score 29; DB 6; Length 1018;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAAAA...AAAAAAAAAAAAA 1478
|||||
DB 987 CCAGTAAAAA...AAAAAAAAAAAAA 1015

```

```

RESULT 23
AX030786
LOCUS       AX030786                1018 bp    DNA          linear      PAT 20-SEP-2000
DEFINITION   Sequence 39 from Patent WO9820023.
ACCESSION   AX030786
VERSION      AX030786.1 GI:10278289
KEYWORDS     .
SOURCE       unidentified.
ORGANISM     unidentified.
REFERENCE    1 (bases 1 to 1018)
AUTHORS      Viney,E.M., Starr,R., Metcalf,D., Nicola,N.A., Willson,T.A.,
              Alexander,W.S., Hilton,D.J., Nicholson,S.E. and Richardson,R.T.
TITLE        Therapeutic and diagnostic agents capable of modulating cellular
              responsiveness to cytokines
JOURNAL      Patent: WO 9820023-A 39 14-MAY-1998;
              INST MEDICAL W & E HALL (AU) ; VINEY ELIZABETH M (AU) ; STARR ROBYN
              (AU) ; METCALF DONALD (AU) ; NICOLA NICOS A (AU) ; WILLSON TRACY A
              (AU) ; ALEXANDER WARREN S (AU) ; HILTON DOUGLAS J (AU) ; NICHOLSON
              SANDRA E (AU) ; RICHARDSON RACHAEL T (AU)
FEATURES     Location/Qualifiers
              source
              1..1018
              /db_xref="taxon:32644"
BASE COUNT   306 a 198 c 211 g 303 t
ORIGIN
Query Match      2.0%; Score 29; DB 6; Length 1018;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
Db 987 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1015

RESULT 24
BD007470
LOCUS       BD007470                1018 bp    DNA          linear      PAT 31-JAN-2002
DEFINITION   Remedies and diagnostics capable of regulating cell response to
              cytokine.
ACCESSION   BD007470
VERSION      BD007470.1 GI:18635841
KEYWORDS     .
SOURCE       unidentified.
ORGANISM     unidentified.
REFERENCE    1 (bases 1 to 1018)
AUTHORS      Hilton,D.J., Alexander,W.S., Viney,E.M., Willson,T.A.,
              Richardson,R.T., Starr,R., Nicholson,S.E., Metcalf,D. and
              Nicola,N.A.
TITLE        Remedies and diagnostics capable of regulating cell response to
              Patent: JP 2001502183-A 28 20-FEB-2001;
              THE WALTER AND ELIZA HALL INSTITUTE OF MEDICAL RESEARCH
JOURNAL      OS Unidentified
              PN JP 2001502183-A/28
              PD 20-FEB-2001
              PF 31-OCT-1997 JP 1998520867
              PR 01-NOV-1996 AU P03384,14-FEB-1997 AU P05117 PI
              DOUGLAS J HILTON,WARREN S ALEXANDER,ELIZABETH M VINEY, PI TRACY A
              WILSON,
              PI RACHAEL T RICHARDSON,ROBYN STARR,SANDRA E NICHOLSON, PI
              DONALD METCALF,
              PI NICOS A NICOLA
              PC C12N15/09,A61K31/7088,A61K38/00,A61K48/00,A61P9/12,A61P35/00,
              PC A61P37/02,
              PC A61P43/00,C07K14/47,C12N15/00,A61K37/02
              CC Strandedness: Single;
              CC Topology: Linear;
              FH Key
              FT source
              1..1018
              /db_xref="taxon:32644"
              /organism="Unidentified".

```

```

FEATURES     Location/Qualifiers
              source
              1..1018
              /db_xref="taxon:32644"
              /organism="unidentified"
BASE COUNT   306 a 198 c 211 g 303 t
ORIGIN
Query Match      2.0%; Score 29; DB 6; Length 1018;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
Db 987 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1015

RESULT 25
AR048809
LOCUS       AR048809                1029 bp    DNA          linear      PAT 29-SEP-1999
DEFINITION   Sequence 6 from patent US 5824306.
ACCESSION   AR048809
VERSION      AR048809.1 GI:6004848
KEYWORDS     .
SOURCE       Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 1029)
AUTHORS      Tang,L. and Blehm,E.Scot.
TITLE        Dirofilaria and Brugia ankyrin proteins, nucleic acid molecules,
              and uses thereof
JOURNAL      Patent: US 5824306-A 6 20-OCT-1998;
              Location/Qualifiers
              source
              1..1029
              /organism="unknown"
BASE COUNT   361 a 173 c 249 g 246 t
ORIGIN
Query Match      2.0%; Score 29; DB 6; Length 1029;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
Db 1001 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1029

RESULT 26
AR048810/c
LOCUS       AR048810                1029 bp    DNA          linear      PAT 29-SEP-1999
DEFINITION   Sequence 8 from patent US 5824306.
ACCESSION   AR048810
VERSION      AR048810.1 GI:6004849
KEYWORDS     .
SOURCE       Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 1029)
AUTHORS      Tang,L. and Blehm,E.Scot.
TITLE        Dirofilaria and Brugia ankyrin proteins, nucleic acid molecules,
              and uses thereof
JOURNAL      Patent: US 5824306-A 8 20-OCT-1998;
              Location/Qualifiers
              source
              1..1029
              /organism="unknown"
BASE COUNT   246 a 249 c 173 g 361 t
ORIGIN
Query Match      2.0%; Score 29; DB 6; Length 1029;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
Db 1001 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1029

RESULT 27
AR048810/c
LOCUS       AR048810                1029 bp    DNA          linear      PAT 29-SEP-1999
DEFINITION   Sequence 8 from patent US 5824306.
ACCESSION   AR048810
VERSION      AR048810.1 GI:6004849
KEYWORDS     .
SOURCE       Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 1029)
AUTHORS      Tang,L. and Blehm,E.Scot.
TITLE        Dirofilaria and Brugia ankyrin proteins, nucleic acid molecules,
              and uses thereof
JOURNAL      Patent: US 5824306-A 8 20-OCT-1998;
              Location/Qualifiers
              source
              1..1029
              /organism="unknown"
BASE COUNT   246 a 249 c 173 g 361 t
ORIGIN
Query Match      2.0%; Score 29; DB 6; Length 1029;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
Db 1001 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1029

```

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Seedi, Jacqueline
 Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Scott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 29 Row: e Column: 7
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 7770224.

FEATURES
 source
 1..1166
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:15929 IMAGE:4559697"
 /tissue_type="kidney, renal cell adenocarcinoma"
 /clone_lib="NIH_MGC_14"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 396..830
 /codon_start=1
 /product="Similar to serine/threonine kinase 10"
 /protein_id="AA014413.1"
 /db_xref="GI:15680140"
 /translation="MASFLFSQFAHLRQSLVYVGSDRGKREAFMILLPCFMPS
 RVLPAHGCECHIRVHTGLLAMPFLPRWGTLLPEVPLSMLGAGVGARAGCW
 HRRHKAQAVLPCCACWPPHRLLCAGDSSAHEASTQSAP"
 BASE COUNT 245 a 319 c 317 g 285 t
 ORIGIN

Query Match 2.0% Score 29; DB 9; Length 1166;
 Best Local Similarity 100.0%; Pred. No. 8.1e-05;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1450 CCAGTAAAAA...AAAAAAAAAAAAA 1478
 |||||...|||||...|||||...|||||...
 Db 1136 CCAGTAAAAA...AAAAAAAAAAAAA 1164

RESULT 31
 HSM802621
 LOCUS
 DEFINITION Homo sapiens mRNA; cDNA DKF2p434I1230 (from clone DKF2p434I1230);
 partial cds.
 ACCESSION AL162083
 VERSION AL162083.1 GI:7328169
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 1243)
 Blum, H., Bauersachs, S., Mewes, H.W., Weil, B. and Wiemann, S.
 Direct Submission
 Submitted (15-MAR-2000) MIPS, Am Klopferspitz 18a, D-82152
 Martinsried, GERMANY
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 Munich/Germany) within the cDNA sequencing consortium of the German
 Genome Project.
 This clone (DKF2p434I1230) is available at the R2PD in Berlin.

Please contact the R2PD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@r2pd.de Further
 information about the clone and the sequencing project is available
 at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES
 source
 1..1243
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="DKF2p434I1230"
 /tissue_type="testis"
 /clone_lib="434 (synonym: htes3). Vector pSport1; host
 DH10B; sites NotI + SalI"
 /dev_stage="adult"
 1..483
 /gene="DKF2p434I1230"
 <1..483
 /gene="DKF2p434I1230"
 /note="Ran binding protein 11 (Homo sapiens)"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CA82416.1"
 /db_xref="GI:7328170"
 /translation="KFGNDOLLGNIEMWVDMNITOPERKLSALALISLPSD,
 NSVQKFKGILINISVEGLHDVMTDPETGYKDCMLSHLEPKVTEDEPTEDQK
 RKRMLAKDPVHTVSLQQFIYEKRAQOEMLGQFQSLMETVDTVEITVLTQLQFLQGF
 "

polya_site 1122
 BASE COUNT 503 a 193 c 215 g 332 t
 ORIGIN

Query Match 2.0% Score 29; DB 9; Length 1243;
 Best Local Similarity 100.0%; Pred. No. 8.1e-05;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1450 CCAGTAAAAA...AAAAAAAAAAAAA 1478
 |||||...|||||...|||||...|||||...
 Db 1121 CCAGTAAAAA...AAAAAAAAAAAAA 1149

RESULT 32
 BC011460
 LOCUS
 DEFINITION Homo sapiens, chitinase 3-like 2, clone MGC:17183 IMAGE:4214121,
 mRNA, complete cds.
 ACCESSION BC011460
 VERSION BC011460.1 GI:15079251
 KEYWORDS MGC.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
 1 (bases 1 to 1449)
 Direct Submission
 Submitted (30-JUL-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center

Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalona@bcm.tmc.edu
 Villalón, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Muzny, D.M., Gibbs, R.A.

```

/product="RV5161"
/protein_id="AAK28356.1"
/db_xref="GI:13506830"
/translation="MSWIFKIQPKQIVSKTVDPDFERYGKLQILEQTKRLQD
MKKSTADLANSSKSAVISQDLLSNPLCEQDDF LHMVYALTDAMKRDAFADNEKYNQ
IQKTVTEPLKKESSIFPSLNARQALQODYGRLOAKVEKYEEKEKTPGLAKLHQ

```

and relationship to other sequences, please visit our Web site (<http://fruitfly.berkeley.edu>) or send email to cdna@fruitfly.berkeley.edu.

FEATURES

source	Location/Qualifiers
1	1601

```

/translation="MSSLORTMLKGLPPTIPGGRIGRADSFKKSRIDRYKPGMWNEF
FAKEDYTVDEORTFLYRKOPKPGPVLILLHGGGYSAITWAFGEVTSMTLHCOC
LCIDMRHGSGKVDDEDDLSADTLAKDIDGLILKLYPEEVOLFVGHSGGATAVHF
AHMALVNLIGITIVDVGTGMEALASQSFRLSRPKYFOSIGNALIMCJIRSGOVNR
VDSAKVSMPOIITNCTNKLATNDPLPDVILEEAHNSMFPNPFISSEDESSPGD
DAADGSSAAAGADKFKNTTKSTTEAKNTWRIDLSKSEKYVWGWSGLSDKFLN
LRLPKQLLASIDGLORTLTIVGQMGREFOMQVLARCGHAVHEDRPHVEAVIISGLIR
NRFAAASEFRCHMPC"
BASE COUNT      419 a  414 c  415 g  353 t
ORIGIN

```

```

Query Match      2.0%; Score 29; DB 3; Length 1601;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAAAA...AAAAAAAAAAAAA 1478
|||||...|||||
DB 1569 CCAGTAAAAA...AAAAAAAAAAAAA 1597

```

```

RESULT 35
BC014817      1642 bp  mRNA  linear  ROD 04-OCT-2001
LOCUS
DEFINITION    Mus musculus, clone MGC:12103 IMAGE:3709357, mRNA, complete cds.
ACCESSION    BC014817
VERSION      BC014817.1 GI:15928705
KEYWORDS     MGC.
SOURCE       house mouse.
ORGANISM     Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 1642)
Strausberg,R.
Direct Submission
Submitted (01-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hqsc.bcm.tmc.edu/cdna/
Contact: villalob@bcm.tmc.edu.
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny,D.M., Gibbs,R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: TRAK Plate: 17 Row: d Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
FEATURES      Location/Qualifiers
1..1642
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="MGC:12103 IMAGE:3709357"
/tissue_type="Mammary tumor. C3(1)-Tag model. Infiltrating
ductal carcinoma, 5 month old virgin mouse."
/clone_lib="MCL_CGAP_Mam6"
/lab_host="Dhi0B"
/note="vector: pCMV-SPORT6"
99..1427
/codon_start=1
/product="Unknown (protein for MGC:12103)"
/protein_id="AAH14817.1"

```

```

CDS
Query Match      2.0%; Score 29; DB 10; Length 1644;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;

```

```

/db_xref="GI:15928706"
/translation="MEESGIETTPPCTPLHPAGLAANVPSTAHAASATSSSPNVSG
MESPPIVSTPQSPGPPVQSPAPPFVSMSPAPSVLSGTSVPPSPSPATAFSGP
PMSHFPPATASGALLUSAPSGPPIISGFSVGTYYDITRHAGRAPOTPLMPFSAPPY
TGILPAPITQOASMTSLAQPGTWSAITPEEQEDPRIINRGODDAPAGGIWGEIKGVA
GNFMVSVLTKRTHSVESMTITLDPGMAPYIKSGGELDTIVTTSNKEVFAAARAFQOE
VFLAVVVGAGQSNIAPOPVGAAGLKGAEIRISLRSGAIEHKQTAVSVENFIAE
LLPKWFDIGCLVVEDPVHGRLEAFTQATPVPLEVQQAOSLT PQDNLNWSGLLVT
VGEVLEKSLNLNVSRTDHLAFTGMSRQMIYSAAKAVAGMYKQRLPPRP"
BASE COUNT      374 a  481 c  435 g  352 t
ORIGIN

```

```

Query Match      2.0%; Score 29; DB 10; Length 1642;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAAAA...AAAAAAAAAAAAA 1478
|||||...|||||
DB 1590 CCAGTAAAAA...AAAAAAAAAAAAA 1618

```

```

RESULT 36
AF327431      1644 bp  mRNA  linear  ROD 05-MAR-2001
LOCUS
DEFINITION    Mus musculus cyclin-dependent kinase 9 (Cdk9), complete cds.
ACCESSION    AF327431
VERSION      AF327431.1 GI:13195442
KEYWORDS     house mouse.
SOURCE       Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 1644)
Bagella,L., MacLachlan,T.K., Buono,R.J., Pisano,M.M., Giordano,A.
and De Luca,A.
Cloning of murine CDK9/PITALRE and its tissue-specific expression
in development
J. Cell. Physiol. 177 (2), 206-213 (1998)
JOURNAL
MEDLINE      98437907
PUBMED       9766517
REFERENCE    2 (bases 1 to 1644)
AUTHORS      Bagella,L., MacLachlan,T.K., Buono,R.J., Pisano,M.M., Giordano,A.
and De Luca,A.
Direct Submission
Submitted (09-DEC-2000) Pathology, Anatomy and Cell Biology, Thomas
Jefferson University, 1020 Locust Street, Philadelphia, PA 19107,
USA
FEATURES      Location/Qualifiers
1..1644
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
1..1644
/gene="Cdk9"
1..1119
/note="PITALRE"
/codon_start=1
/product="cyclin-dependent kinase 9"
/protein_id="AAK15699.1"
/db_xref="GI:13195443"
/translation="MAKQYDSEVCDFECVTEYKELAKIGQCTGFEVEKAKHROTGOK
VALKKVLMENEKGFPTALREIKILQLKHNVNLIETCKRASPYNRCCKSTYLV
DFCEHDLACLLSNLVKFTLSIKRMQMLNGLXYIHRKILLRDMKAAANLITRD
GVKLADFGELARAFSLAKNSQPNRYNRYVTLWYRPELLGERDIYGPIDLWAGCI
MAEWTRSPIMQNGTEOHALISQLCGSITPEVNPVDKYEIFEKLELVKGQRKVK
DLKXAYRDPYALDIDKLVDLPAQIDSDNALNHDFFWSDPMPFSDLKMLSTHLTS
MFEYLAPPRKSGSIQQTQSNRPNATTNQTFFERV"
BASE COUNT      416 a  404 c  455 g  369 t
ORIGIN

```

```

Query Match      2.0%; Score 29; DB 10; Length 1644;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;

```

```

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAAAA... 2.0%; Score 29; DB 9; Length 1665;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1598 CCAGTAAAAA... 2.0%; Score 29; DB 9; Length 1665;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 37
BC002488 Homo sapiens, similar to DKFZP564M2423 protein, clone MGC:1357
LOCUS IMAGE:3051198, mRNA, complete cds.
DEFINITION BC002488.1 GI:12803338
ACCESSION BC002488.1 GI:12803338
VERSION BC002488.1 GI:12803338
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1665)
JOURNAL Strausberg, R.
Direct Submission
Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing by: National Institutes of Health Intramural
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
nisc.mgc@nih.gov
Contact: nisc.mgc@nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
Tiongson, E.F., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: TRAL Plate: 4 Row: a Column: 15
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 5262550.
FEATURES
Source
Location/Qualifiers
1..1665
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:1357 IMAGE:3051198"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH_MGC_15"
/lab_host="DH10B-R"
/notes="vector: pOTB7"
77..1285
/codon_start=1
/product="similar to DKFZP564M2423 protein"
/protein_id="AAH02488.1"
/db_xref="GI:12803339"
/translation="MFGHLQEGFGCVVTRKFDLFDDESDFEVLKAAENKKKAGGG
GVGGPAAKSAQAQATNSNAAGKQVLRKESQKDRKNPLPPSVGVVKKKEETPPVALK
KEGIRRVGRPDQLOGEGLIDREPRPPRPERFEPLBEKGEGEFSDVRPIIDR
PIRGGLGRGGRGMGRGDGFSRGKREFDRHSGDSGLKHEDKRGSGSHNW
GTVKEDELTPIQKQISYNSDLDSNVETTFPGBEHHPPVADHENEVEBEKE
EGPKEMTLDPKAIQNKRAKVFNTRKNEGADGOWKKGFVLHKSSEEAHAEDSYM
DHPKFPANDPITSQLETFNFDLGRPGRGGRGGRGCGRGRNRSRTDKSSASAPDV
DDPEAFPALA"
BASE COUNT 535 a 332 c 458 g 340 t

```

```

ORIGIN
Query Match 2.0%; Score 29; DB 9; Length 1665;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAAAA... 2.0%; Score 29; DB 9; Length 1665;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1621 CCAGTAAAAA... 2.0%; Score 29; DB 9; Length 1665;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 38
AB050426 Macaca fascicularis brain cDNA, clone:QnpA-21709.
LOCUS AB050426
DEFINITION AB050426.1 GI:10998289
ACCESSION AB050426
VERSION AB050426.1 GI:10998289
KEYWORDS fis (full insert sequence).
SOURCE Macaca fascicularis adult male brain parietal lobe cDNA to mRNA,
clone_lib:macaque brain cDNA library QnpA clone:QnpA-21709.
ORGANISM Macaca fascicularis
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Macaca.
REFERENCE 1 (sites)
AUTHORS Osada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirai, M.,
Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.
TITLE Isolation of full-length cDNA clones from macaque brain cDNA
libraries
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1692)
AUTHORS Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.
TITLE Direct Submission
JOURNAL Submitted (21-OCT-2000) to the DDBJ/EMBL/GenBank databases.
Katsuyuki Hashimoto, National Institute of Infectious Diseases,
Division of Genetic Resources, 23-1, Toyama 1-chome, Shinjuku-ku,
Tokyo 162-8640, Japan (E-mail: khashi@nih.go.jp,
URL: http://www.nih.go.jp/yoken/genbank/,
Tel: 81-3-5285-1111 (ex.2120), Fax: 81-3-5285-1181)
Lab host: TOP10
Vector: pME18S-FL3 (Acc.No. AB009864)
R. Site1: DraIII (CACTGTGTC)
R. Site2: DraIII (CACTGTGTC)
Description: 1st strand cDNA was primed with an oligo(dT) primer
[ATGCGCTTTT...TTTTTTTTT]; double-stranded cDNA was synthesized
using specific 5' and 3' primers and amplified by PCR. The PCR
product was digested with SfiI and size selection was performed to
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
the DraIII sites can be used to isolate the cDNA insert. Libraries
were constructed by Sugano et al. (University of Tokyo, Institute of
Medical Science). Custom primer used for sequencing
( 5' end primer [CTTCGCTCTAAAGCTGG];
3' end primer [CGACCTGCGCTCGACCA] ).
FEATURES
Source
Location/Qualifiers
1..1692
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/clone="QnpA-21709"
/sex="male"
/tissue_type="brain parietal lobe"
/clone_lib="macaque brain cDNA library QnpA"
/dev_stage="adult"
BASE COUNT 650 a 249 c 306 g 487 t
ORIGIN
Query Match 2.0%; Score 29; DB 9; Length 1692;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAAAA... 2.0%; Score 29; DB 9; Length 1692;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1659 CCAGTAAAAA... 2.0%; Score 29; DB 9; Length 1692;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 39
BC003901
LOCUS
DEFINITION Mus musculus, Similar to cyclin-dependent kinase 9 (CDC2-related
ACCESSION BC003901
VERSION BC003901.1 GI:13278103
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1710)
Direct Submission
Submitted (28-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC) Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAP Plate: 11 Row: f Column: 18
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

```

FEATURES
    source
    1..1710
        /organism="Mus musculus"
        /db_xref="taxon:10090"
        /clone="MGC:6767 IMAGE:3601310"
        /tissue_type="Mammary tumor. Metallothionien-TGF alpha
        model. 10 month old virgin mouse. Taken by biopsy."
        /clone.lib="NCL_CGAP_Mam1"
        /lab_host="DH10B"
        /note="vector: pCMV-SPORT6"
    74..1192
        /codon_start=1
        /product="Similar to cyclin-dependent kinase 9
        (CDC2-related kinase)"
        /protein_id="AAH03901.1"
        /db_xref="GI:13278104"
        /translation="MAQYDSVECPDCEVTYKELAKIGQTFGEVFKAKHQTGQK
        VALKVLWENKGFPTALREIKLQLKHNVNVLIEICRTKASPNRCCKGSIYLV
        FDFEHLDSLNLVLFKTLSEIKRMQMLNGLYYIHRNKLHRDKMAANVLITRD
        GVLKADFGLARFSLAKNSQPNRYTNRVVLTWYRPPELLLGERDYGPPIDWAGCII
        MAEMWTRSPIMQGNTEQHQLALISLCGSIPTPEVNPVNDKVELFEKLELVKRGKRVK
        DRLKAYVRDPAVLIDKLVLPAQRISDDALNHRDFWSPMPSDLKGLMGLSTHLTS
        MFEVLAPRRKGSQITQOSTNOSRNATTNQIEFEVFF"
BASE COUNT 414 a 423 c 501 g 372 t
ORIGIN
Query Match 2.0%; Score 29; DB 10; Length 1710;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1450 CCAGTAAAAA
Db 1671 CCAGTAAAAA

```

```

RESULT 40
AK021579
LOCUS
DEFINITION Homo sapiens cDNA FLJ11517 fis, clone HEMBA1002337.
ACCESSION AK021579
VERSION AK021579.1 GI:10432787
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens embryo, 10 weeks whole embryo, mainly head cDNA to
mrna, clone.lib.HEMBA1 clone:HEMBA1002337.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (sites)
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosohiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
    Location/Qualifiers
    1..1902
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="HEMBA1002337"
        /clone.lib="HEMBA1"
        /dev_stage="embryo, 10 weeks"
        /tissue_type="whole embryo, mainly head"
        /note="cloning vector: pME18SFL3"
BASE COUNT 605 a 348 c 381 g 568 t
ORIGIN
Query Match 2.0%; Score 29; DB 9; Length 1902;
Best Local Similarity 100.0%; Pred. No. 8.2e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1450 CCAGTAAAAA
Db 1558 CCAGTAAAAA

```

```

RESULT 41
BC009569
LOCUS
DEFINITION Homo sapiens, ankyrin repeat and SOCS box-containing 3, clone
ACCESSION BC009569
VERSION BC009569.1 GI:16306989
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1997)
Strausberg,R.

```

AUTHORS	Welch,M.D., DePace,A.H., Verma,S., Iwamatsu,A. and Mitchison,T.J.
TITLE	The human Arp2/3 complex is composed of evolutionarily conserved subunits and is localized to cellular regions of dynamic actin filament assembly
JOURNAL	J. Cell Biol. 138 (2), 375-384 (1997)
MEDLINE	97375667
PUBMED	9230079
REFERENCE	2 (bases 1 to 2146)
AUTHORS	Welch,M.D., DePace,A.H., Verma,S., Iwamatsu,A. and Mitchison,T.J.
TITLE	Direct Submission
JOURNAL	Submitted (02-JUN-1997) Cellular and Molecular Pharmacology, University of California, San Francisco, 513 Parnassus Ave., San Francisco, CA 94143-0450, USA
FEATURES	Location/Qualifiers
source	1. .2146 /organism="Homo sapiens" /db xref="taxon:9606"

```

gene
1. .2146
/gene="Arp3"
173. .1429
CDS
/gene="Arp3"
/note="one of seven subunits of the Arp2/3 protein
complex; actin-related protein"
/codon_start=1

```

```

/product="Arp3"
/protein_id="AAB64188.1"
/db_xref="GI:2282032"
/translation="WAGRLPACVVDGCGTYKLGAYAGNTEPQFIIPSCIAIKESAKVG
DOQRVMKGVDDLOFFIGDEAIKPTVATKWPIRHGIVEDMDLMEREQVLFYKLYR
AESDEHYLLTEPPLPENRYETAIMEFSFNPVGLYIAVOA/LALAAASWTSGROYCE
RTUJTVDSGGVGHVTPVAGEVYVSGICRKHPIAGRDITYFIQLLRDRRERGVIPEE
OSLETAKAKERYSYCVPLDKFNKDYTDGSKWIKQYTGINAIKKEFSIDVGYERF
LGPEIFHFEANPQTOFISVDEVTQNCIDVRRPLKYNVLISGGSTMERDFGRR
LQDLAKTVDARLKISEELSGRLKPKPIDQVITTHMQRYAVNMFSGSLASLPETYQ
VCHTKDYEIGPSICRHNVPFGVMS"
643 a      387 c      504 g      612 t

BASE COUNT
ORIGIN
2.0% Score 29: DB 9: Length 2146:
Query Match

```

Match: 24/24	Length: 24 bp
Best Local Similarity: 100.0%; Pired No. 8.2e-05;	
Matches: 29; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;	
QY 1450 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478	
DB 2115 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 2143	
RESULT 43	
XLNFAFAS	
LOCUS	
XLNFAFAS	2542 bp
linear	VRT 15-NOV-1994

LOCUS	XLNRAFAS	2542 bp	mRNA	linear	VRT 15-NOV-1994
DEFINITION	X.laevis mRNA for fascin.				
ACCESSION	X75938				
VERSION	X75938.1	GI:436769			
KEYWORDS	fascin.				
SOURCE	African clawed frog.				

ORGANISM
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.

AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE

JOURNAL
submitted (21-DEC-1993) D. KUH, UNIVERSITY OF NIJMEGEN,
TOERNOOIVELD, 6525 ED NIJMEGEN, NETHERLANDS

FEATURES
Location/Qualifiers
1. .2542
Source
/organism="Xenopus laevis"
/db_xref="taxon:8355"

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 10:52:52 ; Search time 176.457 Seconds
(without alignments)
14380.822 Million cell updates/sec

Title: US-09-805-311-7
Perfect score: 1
Sequence: 1 cgaccacgcgtcgggaad.....aaaaaaaaaaaaaaaaaaaaa 1478

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802.*

1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1478	100.0	1478	21 AAA27926	Maize Rad2/FEN-1 c
2	894	60.5	1463	21 AAA27923	Maize Rad2/FEN-1 c
3	850	57.5	1381	21 AAA27925	Maize Rad2/FEN-1 c
4	842	57.0	1541	21 AAA27924	Maize Rad2/FEN-1 c
5	30	2.0	193	22 AAK55174	Human immune/haema
6	30	2.0	818	22 AAD20619	Human ovarian canc
7	30	2.0	851	22 AAS08150	Mammalian toxicolo
8	30	2.0	1454	22 AAS25912	Human cDNA encodin
9	30	2.0	1454	22 AAK64378	Human immune/haema

10	29	2.0	148	22 AAI98818	Human excretory re
11	29	2.0	148	22 AAI63214	Human kidney relat
12	29	2.0	476	22 ABA58117	Human foetal liver
13	29	2.0	476	22 AAK06192	Human brain expres
14	29	2.0	476	22 AAK31850	Human bone marrow
15	29	2.0	476	22 AAH37720	Probe #606 used t
16	29	2.0	646	22 AAH09041	Human cDNA clone (
17	29	2.0	665	22 AAL25133	Human breast cance
18	29	2.0	768	22 AAI97102	Human neuroblastom
19	29	2.0	775	21 AAA61277	Human secreted pro
20	29	2.0	848	22 AAS41014	cDNA encoding nove
21	29	2.0	895	22 AAH34628	Human colon cancer
22	29	2.0	921	19 AAH40521	Homo sapiens CH27
23	29	2.0	1018	19 AAV38864	Homo sapiens SOCS1
24	29	2.0	1029	19 AAV63308	Nucleotide ndiAnk1
25	29	2.0	1029	19 AAV63000	D. immitis ankyrin
26	29	2.0	1029	19 AAV63001	D. immitis ankyrin
27	29	2.0	1029	21 AAA58173	D. immitis ankyrin
28	29	2.0	1029	21 AAA58174	D. immitis ankyrin
29	29	2.0	1306	23 AAS72702	DNA encoding novel
30	29	2.0	1902	22 AAH17037	Human cDNA sequenc
31	29	2.0	2263	21 AAZ36833	cDNA encoding a su
32	29	2.0	4623	22 ABA09048	Human ASB-3 protei
33	29	2.0	5503	19 AAV63314	Full length nucleo
34	29	2.0	5503	19 AAV63020	D. immitis ankyrin
35	29	2.0	5503	19 AAV63021	D. immitis ankyrin
36	29	2.0	5503	21 AAA58193	D. immitis ankyrin
37	29	2.0	5503	21 AAA58194	D. immitis ankyrin
38	28	1.9	54	18 AAT72878	Knot cytochrome P45
39	28	1.9	63	20 AAH16888	Human cancer agent
40	28	1.9	161	22 AAS60388	Cervical cancer pr
41	28	1.9	322	22 AAS07705	Human breast cance
42	28	1.9	330	22 AAL18635	Human breast cance
43	28	1.9	333	22 AAL16837	Human breast cance
44	28	1.9	349	22 AAL26029	Human brain expres
45	28	1.9	382	14 AAO60162	

ALIGNMENTS

RESULT 1
AAA27926
ID AAA27926 standard; cDNA; 1478 BP.
AC AAA27926;
XX
XX
DT 12-SEP-2000 (first entry)
XX
XX
DE Maize Rad2/FEN-1 cDNA.
XX
XX
KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
XX
KW endonuclease; exonuclease; DNA repair; gene targeting; ss.
XX
XX
OS Zea mays.
XX
XX
FH Key Location/Qualifiers
CDS 97..1236
FT /*tag= a
FT
XX
XX
PN WO200036109-A1.
XX
PD 22-JUN-2000.
XX
PF 16-NOV-1999; 99WO-US27147.
XX
XX
PR 15-DEC-1998; 98US-0112332.
XX
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
XX
PI Mahajan PB;
XX
DR WPI; 2000-452026/39.

P-PSDB; AAY95310.

Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA recombination and repair in transgenic plants, e.g. for gene targeting and the production of male sterile plants -

Example 1; Page 77-79; 85pp; English.

The present sequence is that of maize cDNA coding for RAD2/FEN-1 (see AAY95310). The cDNA was isolated from a library prepared from B73 line seed vitreous endosperm RNA. Rad2/FEN-1 is a structure specific endonuclease which under certain conditions also acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to produce the RAD2/FEN-1 polypeptides in transgenic plant cells. The protein is involved in the regulation of DNA repair and recombination in plant systems and therefore may be used for improving gene targeting during further recombinant DNA protocols involving gene targeting. RAD2/FEN-1 endonucleolytic activity is essential in DNA replication and nucleotide excision and repair reactions. The exolytic activity is involved in double strand break repair and joining. The protein is also useful in strand exchange reactions during homologous recombination. These functions may be useful in gene targeting and in the production of male sterile plants. The efficacy of gene targeting can be improved by the overexpression of exogenous Rad2/FEN-1 while male sterile plants can be produced by the down regulation of Rad2/FEN-1 expression.

Sequence 1478 BP; 463 A; 302 C; 365 G; 348 T; 0 other;

Query Match 100.0%; Score 1478; DB 21; Length 1478;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGACCCACCGTCGCGGAAATAGTCGCGGTCGCGGTTCTTGCGCCACTCCGCGTCAGC 60
1 CGACCCACCGTCGCGGAAATAGTCGCGGTCGCGGTTCTTGCGCCACTCCGCGTCAGC 60
61 CGCGCCCGCCACCGCCACAGCCGCCGACGAGATGGGATCGAGGTTTGACGAA 120
61 CGCGCCCGCCACCGCCACAGCCGCCGACGAGATGGGATCGAGGTTTGACGAA 120
121 CTGCTGGCGGACAAATGCGGCCAAGCGGATGAAGGACAGAAATTCGAGAGCTACTTCGCG 180
121 CTGCTGGCGGACAAATGCGGCCAAGCGGATGAAGGACAGAAATTCGAGAGCTACTTCGCG 180
181 CGCAAAATCGCGTCGACGCCAGATGAGATCTACAGTTCCTGATAGTAGTGAAG 240
181 CGCAAAATCGCGTCGACGCCAGATGAGATCTACAGTTCCTGATAGTAGTGAAG 240
241 ACAGGCATGGAACCTCTCACAATCAAGCTGGTGAAGTCACTAGTCTATTTGCAAGGAATG 300
241 ACAGGCATGGAACCTCTCACAATCAAGCTGGTGAAGTCACTAGTCTATTTGCAAGGAATG 300
301 TTCAACCGGACAAATAGATTACTGAAGCGGGAATCAAGCGAGTTAAGTTTGTATGCG 360
301 TTCAACCGGACAAATAGATTACTGAAGCGGGAATCAAGCGAGTTAAGTTTGTATGCG 360
361 AAGCCTCTGATATGAAGAAACAAGAACTTGTCTAAAGATACTCAAAAGAGATGATGCA 420
361 AAGCCTCTGATATGAAGAAACAAGAACTTGTCTAAAGATACTCAAAAGAGATGATGCA 420
421 ACCAAGATCTGACTAGGCGCAGTAGAGTAGAGATAAGATGCGATTGAAAATTTGAGC 480
421 ACCAAGATCTGACTAGGCGCAGTAGAGTAGAGATAAGATGCGATTGAAAATTTGAGC 480
481 AAGAGGACTGTAAGGTCAAGGCAACACACAGCAAGATTGTAACGACTATTAAAGACTT 540
481 AAGAGGACTGTAAGGTCAAGGCAACACACAGCAAGATTGTAACGACTATTAAAGACTT 540
541 ATGGGGTTCCGTGTAGAGCCACTTCTGAACGACAGCAGNATGTCAGCCCTTTCG 600
541 ATGGGGTTCCGTGTAGAGCCACTTCTGAACGACAGCAGNATGTCAGCCCTTTCG 600

601 ATAAACGATAAGGTCTTCGCTGTTGCTTTCAGAGATAAGGACTCCCTTACTTTTGGGGCT 660
601 ATAAACGATAAGGTCTTCGCTGTTGCTTTCAGAGATAAGGACTCCCTTACTTTTGGGGCT 660
661 CCACGGTTCCTTCGTCATTTAATGATCCCAAGTTCCCAAGAAATACCTGTGATGGAATTT 720
661 CCACGGTTCCTTCGTCATTTAATGATCCCAAGTTCCCAAGAAATACCTGTGATGGAATTT 720
721 GATGTTGCCAAGGTTTGGAGGAGCTTGAACCTCAACCTGAGGAGCTTCATTTGTTGTC 780
721 GATGTTGCCAAGGTTTGGAGGAGCTTGAACCTCAACCTGAGGAGCTTCATTTGTTGTC 780
781 ATCCTGTGTTGGATGACATTTATGATAGCATCAAGATATCGGGGGGCAACAGCTCTG 840
781 ATCCTGTGTTGGATGACATTTATGATAGCATCAAGATATCGGGGGGCAACAGCTCTG 840
841 AAATTTATTCGTCACATGGGTCCCATAGAAAGCATCTTGGAGATCTTTAATAAAGACAGA 900
841 AAATTTATTCGTCACATGGGTCCCATAGAAAGCATCTTGGAGATCTTTAATAAAGACAGA 900
901 TATCAAAATTCCTGAGGACTGGCCCTTACCAAGAGCTCGACGCTTGTCAAGGAGCCCTAAT 960
901 TATCAAAATTCCTGAGGACTGGCCCTTACCAAGAGCTCGACGCTTGTCAAGGAGCCCTAAT 960
961 GTCACATTTGGATATTCCTGAGCTAAATGGGACTGACCTGATGAGGAGGCTCTCATAAGT 1020
961 GTCACATTTGGATATTCCTGAGCTAAATGGGACTGACCTGATGAGGAGGCTCTCATAAGT 1020
1021 TTCTCGTGAAGATATGTTTCAATGAAGATCGGTGACAAAGGCCATAGAGAAGATC 1080
1021 TTCTCGTGAAGATATGTTTCAATGAAGATCGGTGACAAAGGCCATAGAGAAGATC 1080
1081 AAATCTGCCAAGAATAAATCGTCGCAAGGAAGACTCGAGTCTCTTTTCAAGCCAACTGCC 1140
1081 AAATCTGCCAAGAATAAATCGTCGCAAGGAAGACTCGAGTCTCTTTTCAAGCCAACTGCC 1140
1141 ACCACATCAGCAGCCGTAACCGGAAGGAGACTTCGGATATAAACAACAGCAGGAGCTGCG 1200
1141 ACCACATCAGCAGCCGTAACCGGAAGGAGACTTCGGATATAAACAACAGCAGGAGCTGCG 1200
1201 AACAAAGAAACAAGGCTGTGGAAGGAAGAAATAATCTTGGATGCTTGTATGACAACTA 1260
1201 AACAAAGAAACAAGGCTGTGGAAGGAAGAAATAATCTTGGATGCTTGTATGACAACTA 1260
1261 CGACTACGAAGCAGCGGTGCATGATCAGTTCGCTAGATTTAATTAACCTCCCTGTTTA 1320
1261 CGACTACGAAGCAGCGGTGCATGATCAGTTCGCTAGATTTAATTAACCTCCCTGTTTA 1320
1321 ACTCAGACCTTTGGTGAAGTTTGGCCCATGTTTCAAGCTGGGGTAAGTTAGTTGTGTTG 1380
1321 ACTCAGACCTTTGGTGAAGTTTGGCCCATGTTTCAAGCTGGGGTAAGTTAGTTGTGTTG 1380
1381 AAGAGATTGGTGTACCAAGTAAACAACTATCGCTGTTTTTACTTCTGCTCTTGAA 1440
1381 AAGAGATTGGTGTACCAAGTAAACAACTATCGCTGTTTTTACTTCTGCTCTTGAA 1440
1441 GTATGATGCCAGTAAACAAACAAAAA 1478
1441 GTATGATGCCAGTAAACAAACAAAAA 1478

RESULT 2

AAA27923

ID AAA27923 standard; cDNA; 1463 BP.

XX AAA27923;

XX AC AC

XX 12-SEP-2000 (first entry)

XX Maize Rad2/FEN-1 cDNA.

XX Maize; Rad2/FEN-1; transgenic plant; male sterile plant;

KW endonuclease; exonuclease; DNA repair; gene targeting; ss.

DE Maize Rad2/FEN-1 cDNA.
 XX Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
 KW endonuclease; exonuclease; DNA repair; gene targeting; ss.
 XX Zea mays.
 FH Key Location/Qualifiers
 FT CDS 37..1176
 FT /*tag= a
 XX
 PN WO200036109-A1.
 XX 22-JUN-2000.
 PD
 XX 16-NOV-1999; 99WO-US27147.
 PF
 XX 15-DEC-1998; 98US-0112332.
 PR
 XX (PTON-) PIONEER HI-BRED INT INC.
 PA
 XX MahaJan PB;
 PI
 XX WPI: 2000-452026/39.
 DR P-PSDB; AAY95309.
 XX
 XX Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
 PT recombination and repair in transgenic plants, e.g. for gene targeting
 PT and the production of male sterile plants -
 XX Example 1; Page 75-76; 85pp; English.
 XX
 XX The present sequence is that of maize cDNA coding for RAD2/FEN-1
 CC (see AAY95309). The cDNA was isolated from a cDNA library prepared
 CC from W23 line tassel (16-18 cm long) polyA RNA. Rad2/FEN-1 is a
 CC structure specific endonuclease which under certain conditions also
 CC acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to
 CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
 CC The protein is involved in the regulation of DNA repair and
 CC recombination in plant systems and therefore may be used for
 CC improving gene targeting during further recombinant DNA protocols
 CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential
 CC in DNA replication and nucleotide excision and repair reactions.
 CC The exolytic activity is involved in double strand break repair and
 CC end joining. The protein is also useful in strand exchange
 CC reactions during homologous recombination. These functions may be
 CC useful in gene targeting and in the production of male sterile
 CC plants. The efficacy of gene targeting can be improved by the
 CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
 CC be produced by the down regulation of Rad2/FEN-1 expression.
 XX
 SQ Sequence 1381 BP; 441 A; 269 C; 346 G; 325 T; 0 other;
 Query Match 57.5%; Score 850; DB 21; Length 1381;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1200; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 76 GCCACAGCCCGCAGACGATGGGCATCAAGGGTTTGACGAAACTGCTGGCGGCAAT 135
 DB 16 GCCACAGCCCGCAGACGATGGGCATCAAGGGTTTGACGAAACTGCTGGCGGCAAT 75
 QY 136 GCGCCCAAGCGGATGAAGAGCAGAGTTGCGAGCTACTTGGCGCGCAAAATCGCCGTC 195
 DB 76 GCGCCCAAGCGGATGAAGAGCAGAGTTGCGAGCTACTTGGCGCGCAAAATCGCCGTC 135
 QY 196 GACGCCAGCATGAGCATCTACCAAGTTCCTGATAGTAGTTGGAAGCAGAGGATGGAAC 255
 DB 136 GACGCCAGCATGAGCATATACCAAGTTCCTGATTTGTTGGAAGCAGAGGATGGAAC 195
 QY 256 CTCACAAATGAAGCTGGTGAAGTCTAGTCAATTTTTCGAAGGAATTTCAACCGGACAATA 315
 DB 196 CTCACAAATGAAGCTGGTGAAGTCTAGTCAATTTTTCGAAGGAATTTTCACCGGACAATA 255

QY 316 AGATTACTGGAAGCGGAATCAAGCCAGTTTATGTTTGTGATGCAAGCCTCCTGATATG 375
 DB 256 AGATTACTGGAAGCGGAATCAAGCCAGTTTATGTTTGTGATGCAAGCCTCCTGATATG 315
 QY 376 AAGAAACAAGAACTTGCTAAAAGATCTCAAAAAGAGATGATGCAACAAGATCTGACT 435
 DB 316 AAGAAACAAGAGCTTGCTAAAAGATCTCAAAAAGAGATGATGCAACAAGATCTGACT 375
 QY 436 GAGCAGTAGAGGTAGGATGAAGATGCGATTGAAAAATTTGAGCAAGAGGACTGTAAG 495
 DB 376 GAGCAGTAGAGGTAGGATGAAGATGCGATTGAAAAATTTGAGCAAGAGGACTGTAAG 435
 QY 496 GTCACAAGGCAACACGAAGATTGTAACGACTATTAAAGACTTATGGGGGTTCCCTGTT 555
 DB 436 GTCACAAGGCAACACGAAGATTGTAACGCGCTATTAAAGACTTATGGGGGTTCCCTGTT 495
 QY 556 GTAGAGGCACCTTCTGAAGCAGAGCAGAAATGTGCAGCCCTTTGCATATAAGAGGTG 615
 DB 496 GTAGAGGCACCTTCTGAAGCAGAGCAGAAATGTGCAGCCCTTTGCATATAAGAGGTG 555
 QY 616 TTGCTGTTGCTTTCAGAAGATGAAGACTCCCTTACTTTTGGGGTCCACGGTTCCTTCGT 675
 DB 556 TTGCTGTTGCTTTCAGAAGATGAAGACTCCCTTACTTTTGGGGTCCACGGTTCCTTCGT 615
 QY 676 CATTTAATGATCCCAAGTTCCCAAGAAATACCTGTGATGGAATTTGATGTGCCAAGTT 735
 DB 616 CATTTAATGATCCCAAGTTCCCAAGAAATACCTGTGATGGAATTTGATGTGCCAAGTT 675
 QY 736 TTGAGGAGGCTTCAACTCACCATGACGAGTTCATTGATTTGTGCATCCTGTGGAATGT 795
 DB 676 TTGAGGAGGCTTCAACTCACCATGACGAGTTCATTGATTTGTGCATCCTGTGGAATGT 735
 QY 796 GACTATTGTGATAGCATCAAGTATCGGGGGCAACAGCTCTGAAACTTATTTCGTCAA 855
 DB 736 GACTATTGTGATAGCATCAAGTATCGGGGGCAACAGCTCTGAAACTTATTTCGTCAA 795
 QY 856 CATGGGTCCATGAAAGCATCTTGGAGAACTCTTAATAAGACAGATATCAAAATTCCTGAG 915
 DB 796 CATGGGTCCATGAAAGCATCTTGGAGAACTCTTAATAAGACAGATATCAAAATTCCTGAG 855
 QY 916 GACTGGGCTTTACCAAGAAGCTCGAGCTTGTTCAGAGGAGCCTTAATGTGCATTTGGATATT 975
 DB 856 GACTGGGCTTTACCAAGAAGCTCGAGCTTGTTCAGAGGAGCCTTAATGTGCATTTGGATATT 915
 QY 976 CTTGAGTAAATGGAGTGCACCTGATGAGAGGCTCTCATAACTTTCCTGGTAAAGAT 1035
 DB 916 CTTGAGTAAATGGAGTGCACCTGATGAGAGGCTCTCATAACTTTCCTGGTAAAGAT 975
 QY 1036 AATGGTTTCAATGAAGATCGGTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAAT 1095
 DB 976 AATGGTTTCAATGAAGATCGGTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAAT 1035
 QY 1096 AATTCGTGCGAAGAGACTCGAGTCCCTTTTCAAGCCAACTGCCACATCAGCACCG 1155
 DB 1036 AATTCGTGCGAAGAGACTCGAGTCCCTTTTCAAGCCAACTGCCACATCAGCACCG 1095
 QY 1156 CTAACCGGAAGAGACTTCGGATAAACAAGAGGAGCTGCGCAACAGAAACAAAG 1215
 DB 1096 CTAACCGGAAGAGACTTCGGATAAACAAGAGGAGCTGCGCAACAGAAACAAAG 1155
 QY 1216 GCTGCTGGAAGAGAAATATCTTGGATGCTTGTATGTAACACTACGACTACGAAAGCAG 1275
 DB 1156 GCTGCTGGAAGAGAAATATCTTGGATGCTTGTATGTAACACTACGACTACGAAAGCAG 1215
 QY 1276 CGGTGGC 1282
 DB 1216 CGGTGGC 1222

RESULT 4
 AAA27924
 ID AAA27924 standard; cdNA; 1541 BP.
 XX


```
RESULT 5
AAK55174
ID AAK55174 standard; cDNA; 193 BP.
XX
AC AAK55174;
XX
DT 06-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:234.
DE
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytotaxitic; gene therapy; vaccine; metastasis; ss.
KW
XX Homo sapiens.
XX
XX WO200157182-A2.
PN
PN 09-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0215647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234397.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0245474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0248207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
```

```

PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-483426/52.
XX P-PSDB; AAM82393.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX Claim 1; SEQ ID NO 234; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention.
XX
XX Sequence 193 BP; 91 A; 31 C; 34 G; 36 T; 1 other;
XX
XX Query Match 2.0%; Score 30; DB 22; Length 193;
XX Best Local Similarity 100.0%; Pred. No. 0.012;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1449 GCCAGTAAACAAAAAAGAAAAAAGAAAAA 1478
XX ||||||
XX Db 139 GCCAGTAAACAAAAAAGAAAAAAGAAAAA 168
XX
XX RESULT 6
XX AAD20619
XX ID AAD20619 standard; cDNA; 818 BP.
XX
XX AC AAD20619;
XX
XX DT 03-JAN-2002 (first entry)
XX
XX DE Human ovarian cancer antigen-encoding gene 2 cDNA clone HUSYJ75.
XX
XX KW Human; ovarian cancer antigen; proliferative disorder; cancer; tumour;
XX respiratory system disorder; asthma; hematopoietic disorder; skin aging;
XX immune system disorder; AIDS; skin disorder; autoimmune disease; allergy;
XX rheumatoid arthritis; inflammation; neurological disorder; septic shock;
XX Alzheimer's disease; Parkinson's disease; diabetes; angiogenic disorder;
XX atherosclerosis; cardiovascular disorder; gastrointestinal disorder;
XX epithelial cell proliferation; transplantation; chemotaxis; infection;
XX food additive; wound healing; endocrine disorder; kidney disorder;

```

```

KW gene therapy; cytostatic; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 155..499
XX FT /*tag= a
XX FT /product= "Human ovarian cancer antigen"
XX FT 155..187
XX FT /*tag= b
XX FT 188..496
XX FT /*tag= c
XX
XX WO200170804-A1.
XX
XX 27-SEP-2001.
XX
XX 16-MAR-2001; 2001WO-US08585.
XX
XX 17-MAR-2000; 2000US-190076P.
XX 23-AUG-2000; 2000US-227009P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
XX
XX WPI; 2001-639119/73.
XX P-PSDB; AAE12624.
XX
XX Nucleic acids encoding 7 ovarian cancer antigen polypeptides, useful
XX for preventing, diagnosing and/or treating ovarian cancer and ovarian
XX cancer metastases -
XX
XX Claim 1; Page 408; 427pp; English.
XX
XX AAD20618-AAD20624 represent cDNAs corresponding to 7 human ovarian
XX and ovarian cancer associated protein (collectively known as ovarian
XX cancer antigens) genes, and AAE12623-AAE12629 represent the proteins
XX they encode. AAE12630-AAE12638 represent human ovarian cancer antigenic
XX fragments or variants. Ovarian cancer antigens and their corresponding
XX DNAs are used in the prevention, diagnosis and treatment of diseases
XX associated with their inappropriate expression. These disorders include
XX proliferative disorders, cancer, tumours, respiratory system disorders,
XX asthma, hematopoietic disorders, diseases of the immune system, AIDS,
XX skin disorders, autoimmune diseases (e.g., rheumatoid arthritis),
XX inflammation, allergies, neurological disorders (e.g., Alzheimer's
XX disease, Parkinson's disease), septic shock, diabetes, atherosclerosis,
XX cardiovascular disorders, angiogenic disorders, kidney disorders,
XX gastrointestinal disorders, endocrine disorders and infections. The
XX proteins can also be used to aid wound healing and epithelial cell
XX proliferation, to prevent skin aging due to sunburn, to maintain organs
XX before transplantation, to identify their cognate ligands or binding
XX partners, in chemotaxis and can be used as a food additive. Antibodies
XX specific for a protein of the invention can be used in alleviating
XX symptoms associated with the disorders mentioned above and in diagnostic
XX immunoassays e.g., enzyme linked immunosorbent assay (ELISA). The DNA of
XX the invention is used in gene therapy. The present sequence represents
XX a human ovarian cancer antigen-encoding cDNA of the invention.
XX
XX Sequence 818 BP; 172 A; 334 C; 199 G; 113 T; 0 other;
XX
XX Query Match 2.0%; Score 30; DB 22; Length 818;
XX Best Local Similarity 100.0%; Pred. No. 0.01;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1449 GCCAGTAAACAAAAAAGAAAAAAGAAAAA 1478
XX ||||||
XX Db 778 GCCAGTAAACAAAAAAGAAAAAAGAAAAA 807
XX
XX RESULT 7
XX AAS08150

```



```
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
```

```
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-488783/53.
XX P-PSDB; AAU15925.
XX
XX New nucleic acid molecules encoding 461 human secreted proteins for
PT diagnosing, preventing, treating or ameliorating medical conditions and
PT used as food additives or preservatives -
XX
XX Claim 1; SEQ ID NO 91; 980pp; English.
XX
XX The invention relates to isolated nucleic acid molecules and their
CC encoded secreted proteins. The nucleic acids and proteins are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
CC are also used in diagnosing a pathological condition or susceptibility
CC to a pathological condition. Antibodies to the proteins can also
CC be used in alleviating symptoms associated with the disorders and in
CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
CC include autoimmune diseases e.g. rheumatoid arthritis,
CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi
CC and ocular disorders e.g. corneal infection, and many other
CC disorders listed in the specification. The polypeptides can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The polypeptides can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, cofactors and other nutritional components. The present
CC sequence encodes a novel secreted protein of the invention.
XX
XX Query Match 2.0%; Score 30; DB 22; Length 1454;
XX Best Local Similarity 100.0%; Pred. No. 0.0093;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1449 GCCAGTAAAAAATAAAAAAATAAAAAA 1478
Db 1411 GCCAGTAAAAAATAAAAAAATAAAAAA 1440
XX
XX RESULT 9
XX AAK64378
XX ID AAK64378 standard; cDNA; 1454 BP.
XX AC AAK64378;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:9438.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ss.
XX
XX Homo sapiens.
XX
```

PR WO200157182-A2.
XX 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US01354.
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225256.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227183.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231124.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232197.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX
PI

```
XX WPI; 2001-483426/52.
DR P-PSDB; AAM91597.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis.
XX
PS Claim 1: SEQ ID NO 9439; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 1454 BP; 294 A; 475 C; 384 G; 301 T; 0 other;

Query Match          2.08; Score 30; DB 22; Length 1454;
Best Local Similarity 100.0%; Pred. No. 0.0093;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1449 GCACGTAAAAAATAAAAAAAAAAAAAAAAAA 1478
Db 1411 GCCAGTAAAAAATAAAAAAAAAAAAAAAAAA 1440

RESULT 10
AA198818
ID AA198818 standard; cDNA; 148 BP.
XX
AC AA198818;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human excretory related polynucleotide SEQ ID NO 262.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulneryary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine;
KW excretory system; ss.
XX
OS Homo sapiens.
XX
PN WO20015313-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01323.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 05-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 06-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 08-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0233223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
```

PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244517.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PR Rosen CA, Barash SC, Ruben SM;
 PR WPI; 2001-465569/50.
 PR P-PSDB; AAM99845.
 PR Isolated nucleic acid molecule encoding excretory system antigen is
 PR used in preventing, treating or ameliorating a medical condition -
 PR
 PR Claim 1; SEQ ID NO 262; 574pp + Sequence Listing: English.
 PR
 PR The invention relates to novel excretory system related human
 PR polynucleotides (AA198567-AA199503) and the encoded proteins
 PR (AAM99594-AAM99913) useful for preventing, treating or ameliorating
 PR medical conditions e.g. by protein or gene therapy, especially

CC disorders related to the excretory system. The genes are isolated
 CC from a range of human tissues disclosed in the specification. The
 CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 148 BP; 69 A; 21 C; 21 G; 37 T; 0 other;

Query Match 2.0%; Score 29; DB 22; Length 148;
 Best Local Similarity 100.0%; Pred. No. 0.033;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1450 CCAGTAAAAA 1478
 Db 110 CCAGTAAAAA 138

RESULT 11
 AA163214
 ID AA163214 standard; cDNA; 148 BP.
 XX
 AC AAT63214;
 DT 22-OCT-2001 (first entry)
 DE Human kidney related polynucleotide SEQ ID NO 254.
 KW Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;
 KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
 KW neotropic; neuroprotective; antibacterial; virucide; fungicide;
 KW ophthalmological; antiallergic; hepatotropic; antidiabetic;
 KW antiinflammatory; antitumor; cancer; immune disorder; antiparasitic;
 KW gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; ss.
 XX Homo sapiens.
 OS
 XX
 PN W020015323-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01343.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228944.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235835.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 29-SEP-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-488784/53.
DR P-PSDB; AAM42660.
DR
DR
XX
PT New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating and/or preventing human diseases and disorders -
XX
PS Claim 1; SEQ ID NO 254; 564pp + Sequence Listing; English.
XX
CC The invention relates to novel kidney related polynucleotides
CC (AA163791-AA163793) and the encoded polypeptides (AAM42417-AAM42691)
CC collectively known as kidney antigens and the use of such kidney antigens
CC for detecting disorders of the kidney, especially kidney cancer and
CC kidney cancer metastases. The polynucleotides and proteins are also
CC useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. The genes are isolated from a range
CC of human tissues disclosed in the specification. The nucleic acids,
CC proteins, antibodies and (ant)agonists are useful in the diagnosis,
CC treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer,
CC and other cancers of the adrenal gland, bone, bone marrow, breast,
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders
CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
CC sclerosis, rheumatoid arthritis and ulcerative colitis;
CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound

CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
CC and (f) infectious diseases such as viral, bacterial, fungal and
CC parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 148 BP; 69 A; 21 C; 21 G; 37 T; 0 other;

Query Match 2.0%; Score 29; DB 22; Length 148;
Best Local Similarity 100.0%; Pred. No. 0.033;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1450 CCAGTAAAAA 1478

Db 110 CCAGTAAAAA 138

RESULT 12

ABA58117

ID ABA58117 standard; DNA; 476 BP.

XX

AC ABA58117;

XX

DT 01-FEB-2002 (first entry)

XX

DE Human foetal liver single exon nucleic acid probe #6422.

XX

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX

OS Homo sapiens.

XX

PN WO200157277-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US00669.

XX

PR 04-FEB-2000; 2000US-0180312.

XX

PR 26-MAY-2000; 2000US-0207456.

XX

PR 30-JUN-2000; 2000US-0608408.

XX

PR 03-AUG-2000; 2000US-0632366.

XX

PR 21-SEP-2000; 2000US-0234687.

XX

PR 27-SEP-2000; 2000US-0236359.

XX

PR 04-OCT-2000; 2000GB-0024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

XX WPI; 2001-483447/52.

XX

PT Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human foetal liver -

XX

PS Claim 1; SEQ ID NO 6422; 639pp + sequence listing; English.

XX

CC The invention relates to a single exon nucleic acid probe for

XX measuring human gene expression in a sample derived from human foetal

XX liver. The single exon nucleic acid probes may be used for predicting,

XX measuring and displaying gene expression in samples derived from human

XX foetal liver. The present sequence is a single exon nucleic acid

XX probe of the invention.

XX Note: The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 476 BP; 170 A; 107 C; 70 G; 129 T; 0 other;

XX

Query Match 2.0%; Score 29; DB 22; Length 476;

Best Local Similarity 100.0%; Pred. No. 0.028;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1450 CCAGTAAAAA 1478

Db 34 CCAGTAAAAA 62

RESULT 13

AAK06192

ID AAK06192 standard; DNA; 476 BP.

XX

AC AAK06192;

XX

DT 05-NOV-2001 (first entry)

XX

DE Human brain expressed single exon probe SEQ ID NO: 6183.

XX

KW Human; brain expressed exon; gene expression analysis; probe;

XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

XX epilepsy; cancer; ss.

XX

OS Homo sapiens.

XX

PN WO200157275-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US00667.

XX

PR 04-FEB-2000; 2000US-0180312.

XX

PR 26-MAY-2000; 2000US-0207456.

XX

PR 30-JUN-2000; 2000US-0608408.

XX

PR 03-AUG-2000; 2000US-0632366.

XX

PR 21-SEP-2000; 2000US-0234687.

XX

PR 27-SEP-2000; 2000US-0236359.

XX

PR 04-OCT-2000; 2000GB-0024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

XX WPI; 2001-483446/52.

XX

PT Single exon nucleic acid probes for analyzing gene expression in human

XX brains -

XX

PS Example 4; SEQ ID NO: 6183; 650pp + Sequence Listing; English.

XX

CC The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX brain. They can be used to measure gene expression in brain cell samples,

XX which may enable the diagnosis and improved treatment of nervous system

XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

XX epilepsy and cancers. The present sequence is one of the probes of the

XX invention.

XX Sequence 476 BP; 170 A; 107 C; 70 G; 129 T; 0 other;

XX

Query Match 2.0%; Score 29; DB 22; Length 476;

Best Local Similarity 100.0%; Pred. No. 0.028;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1450 CCAGTAAAAA 1478

Db 34 CCAGTAAAAA 62

RESULT 14

AAK31850

ID AAK31850 standard; DNA; 476 BP.

XX

AC AAK31850;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human bone marrow expressed single exon probe SEQ ID NO: 6407.
 XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00668.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 6407; 658pp + Sequence Listing; English.
 XX
 CC: The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention.
 XX
 SQ Sequence 476 BP; 170 A; 107 C; 70 G; 129 T; 0 other;
 XX
 Query Match 2.0%; Score 29; DB 22; Length 476;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 25; SEQ ID NO 6406; 654pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 XX
 SQ Sequence 476 BP; 170 A; 107 C; 70 G; 129 T; 0 other;
 XX
 Query Match 2.0%; Score 29; DB 22; Length 476;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1450 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
 DB 34 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 62
 XX
 RESULT 15
 ID AAI37720 standard; DNA; 476 BP.
 XX
 AC AAI37720;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Probe #6406 used to measure gene expression in human placenta sample.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157272-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00663.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 XX

PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488997/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 25; SEQ ID NO 6406; 654pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 XX
 SQ Sequence 476 BP; 170 A; 107 C; 70 G; 129 T; 0 other;
 XX
 Query Match 2.0%; Score 29; DB 22; Length 476;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1450 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
 DB 34 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 62
 XX
 RESULT 16
 ID AAH09041/c
 ID AAH09041 standard; cDNA; 646 BP.
 XX
 AC AAH09041;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA clone (3'-primer) SEQ ID NO:5876.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 3; SEQ ID 5876; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 646 BP; 206 A; 117 C; 97 G; 216 T; 10 other;

Query Match 2.0%; Score 29; DB 22; Length 646;
Best Local Similarity 100.0%; Pred. No. 0.027; Mismatches 0; Indels 0; Gaps 0;
Matches 29; Conservative 0;

Qy 1450 CCAGTAAAGAAAAAAGAAAAAAGAAAAA 1478
|||||
Db 345 CCAGTAAAGAAAAAAGAAAAAAGAAAAA 317

RESULT 17
AAL25133/c
ID AAL25133 standard; cDNA; 665 BP.

XX AAL25133;
XX 07-DEC-2001 (first entry)
XX Human breast cancer expressed polynucleotide 17590.
XX Human; breast cancer; cell marker; cytostatic; ss.
XX Homo sapiens.
XX WO200151628-A2.
XX 19-JUL-2001.
XX 10-JAN-2001; 2001WO-US00798.
XX 14-JAN-2000; 2000US-0176077.
XX 14-MAR-2000; 2000US-0189167.
XX 24-MAR-2000; 2000US-0192099.
XX 29-MAR-2000; 2000US-0193480.
XX 15-MAY-2000; 2000US-0205230.
XX 09-JUN-2000; 2000US-0211315.
XX 25-JUL-2000; 2000US-0220534.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2001-451856/48.
XX New peptide useful as a marker for the diagnosis of breast cancer
XX Claim 1; Page 3254; 3695pp; English.

CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.

XX Sequence 665 BP; 163 A; 125 C; 118 G; 259 T; 0 other;

Query Match 2.0%; Score 29; DB 22; Length 665;
Best Local Similarity 100.0%; Pred. No. 0.026; Mismatches 0; Indels 0; Gaps 0;
Matches 29; Conservative 0;

Qy 1450 CCAGTAAAGAAAAAAGAAAAAAGAAAAA 1478
|||||
Db 38 CCAGTAAAGAAAAAAGAAAAAAGAAAAA 10

RESULT 18
AAI97102/c
ID AAI97102 standard; cDNA; 768 BP.

XX AAI97102;
XX 13-NOV-2001 (first entry)
XX Human neuroblastoma expressed polynucleotide SEQ ID NO 3177.
XX Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX Homo sapiens.
XX WO200166719-A1.
XX 13-SEP-2001.
XX 02-MAR-2001; 2001WO-JP01629.
XX 07-MAR-2000; 2000JP-0159195.
XX (CHIB-) CHIBA PREFECTURE.
XX (HISM) HISAMITSU PHARM CO LTD.
XX Nakagawara A;
XX WPI; 2001-565584/63.
XX Nucleic acids originating in gene expressed in human neuroblastoma,
XX useful as probe or primer in diagnosing prognosis of human
XX neuroblastoma, malignancy and susceptibility indicator or tumour marker
XX for anti-cancer agents -
XX Claim 1; Page 2311; 2979pp; Japanese.
XX The invention relates to novel genes (AAI93926-AAI97963) expressed in
XX human neuroblastoma. The nucleic acids are applicable as a probe or
XX primer in diagnosing the prognosis of human neuroblastoma, malignancy and
XX susceptibility indicators or tumour markers for anti-cancer agents. The
XX gene information for diagnosing prognosis is related to factors similar
XX to that for N-myc and TrkA genes.

XX Sequence 768 BP; 262 A; 106 C; 119 G; 241 T; 40 other;

Query Match 2.0%; Score 29; DB 22; Length 768;
Best Local Similarity 100.0%; Pred. No. 0.026; Mismatches 0; Indels 0; Gaps 0;
Matches 29; Conservative 0;

Qy 1450 CCAGTAAAGAAAAAAGAAAAAAGAAAAA 1478
|||||
Db 75 CCAGTAAAGAAAAAAGAAAAAAGAAAAA 47


```

RESULT 19
AAA61277
ID AAA61277 standard; DNA: 775 BP.
AC
XX
AC AAA61277;
XX
DT 18-OCT-2000 (first entry)
XX
DE Human secreted protein gene 18 clone HE2FL70.
XX
KW Human; secreted protein; fusion protein; gene therapy;
KW protein therapy; diagnosis; tissue; cancer; tumour; AIDS;
KW autoimmune disorder; allergy; cardiovascular; viral; bacterial;
KW fungal infection; immunosuppressive; ds.
XX
OS Homo sapiens.
XX
PN WO200029422-A1.
XX
XX 25-MAY-2000.
XX
PF 09-NOV-1999; 99WO-US26409.
XX
PR 12-NOV-1998; 98US-0108207.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ni J, Ruben SM, Rosen CA, Ebner R, Florence KA, Young PE;
PI Birse CE, Carter KC, Komatsoulis G;
XX
XX WPI: 2000-387729/33.
XX
PT Novel human secreted proteins useful for diagnosing, preventing,
PT treating and ameliorating a medical condition e.g. cardiovascular
PT disease -
XX
XX Claim 1; Page 240; 295pp; English.
XX
CC The present sequence represents a nucleic acid molecule which encodes a
CC secreted human protein. The gene number and the clone it was derived
CC from are given in the descriptor line.
CC The invention relates to 31 novel genes and their fragments (nucleic
CC acid sequences: AAA61260-A61293; amino acid sequences AAB12301-B12371)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new
CC polynucleotides. Specific uses are described for each of the 31
CC and include products for the diagnosis or treatment of cancer, tumours,
CC AIDS, autoimmune disorders, allergy, cardiovascular disorders, viral
CC bacterial and fungal infection. The genes are used to generate fusion
CC proteins by linking to the gene a human immunoglobulin portion (AAA61251)
CC for increasing stability of the fused protein as compared to the
CC secreted protein only.
XX
SQ Sequence 775 BP; 212 A; 176 C; 158 G; 229 T; 0 other;
Query Match 2.0%; Score 29; DB 21; Length 775;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1450 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
Db 709 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 737
RESULT 20
AAS41014
ID AAS41014 standard; CDNA: 848 BP.
XX

```

```

AC AAS41014;
XX
XX 17-DEC-2001 (first entry)
XX
DE CDNA encoding novel human enzyme polypeptide #230.
XX
XX
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.
XX
OS Homo sapiens.
XX
PN WO200155301-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01239.
XX
XX 31-JAN-2000; 2000US-0179065.
XX
PR 04-FEB-2000; 2000US-0180628.
XX
PR 24-FEB-2000; 2000US-0184654.
XX
PR 02-MAR-2000; 2000US-0186350.
XX
PR 16-MAR-2000; 2000US-0189874.
XX
PR 17-MAR-2000; 2000US-0190076.
XX
PR 18-APR-2000; 2000US-0198123.
XX
PR 19-MAY-2000; 2000US-0205515.
XX
PR 07-JUN-2000; 2000US-0209467.
XX
PR 28-JUN-2000; 2000US-0214886.
XX
PR 30-JUN-2000; 2000US-0215135.
XX
PR 07-JUL-2000; 2000US-0216647.
XX
PR 07-JUL-2000; 2000US-0216880.
XX
PR 11-JUL-2000; 2000US-0217487.
XX
PR 11-JUL-2000; 2000US-0217496.
XX
PR 14-JUL-2000; 2000US-0218290.
XX
PR 26-JUL-2000; 2000US-0220963.
XX
PR 14-AUG-2000; 2000US-0220964.
XX
PR 14-AUG-2000; 2000US-0224518.
XX
PR 14-AUG-2000; 2000US-0224519.
XX
PR 14-AUG-2000; 2000US-0225213.
XX
PR 14-AUG-2000; 2000US-0225266.
XX
PR 14-AUG-2000; 2000US-0225267.
XX
PR 14-AUG-2000; 2000US-0225268.
XX
PR 14-AUG-2000; 2000US-0225270.
XX
PR 14-AUG-2000; 2000US-0225271.
XX
PR 14-AUG-2000; 2000US-0225275.
XX
PR 14-AUG-2000; 2000US-0225758.
XX
PR 14-AUG-2000; 2000US-0225759.
XX
PR 18-AUG-2000; 2000US-0226279.
XX
PR 22-AUG-2000; 2000US-0226681.
XX
PR 22-AUG-2000; 2000US-0226868.
XX
PR 22-AUG-2000; 2000US-0227182.
XX
PR 23-AUG-2000; 2000US-0227009.
XX
PR 30-AUG-2000; 2000US-0228924.
XX
PR 01-SEP-2000; 2000US-0229287.
XX
PR 01-SEP-2000; 2000US-0229343.
XX
PR 01-SEP-2000; 2000US-0229344.
XX
PR 01-SEP-2000; 2000US-0229345.
XX
PR 05-SEP-2000; 2000US-0229509.
XX
PR 05-SEP-2000; 2000US-0229513.
XX
PR 06-SEP-2000; 2000US-0230437.
XX
PR 06-SEP-2000; 2000US-0230438.
XX
PR 08-SEP-2000; 2000US-0231242.
XX
PR 08-SEP-2000; 2000US-0231243.
XX
PR 08-SEP-2000; 2000US-0231244.
XX
PR 08-SEP-2000; 2000US-0231413.
XX
PR 08-SEP-2000; 2000US-0231414.
XX
PR 08-SEP-2000; 2000US-0232080.
XX
PR 08-SEP-2000; 2000US-0232081.
XX
PR 12-SEP-2000; 2000US-0231968.

```



```

XX PD 05-APR-2001.
XX PF 28-SEP-2000; 2000WO-US26524.
XX PR 29-SEP-1999; 99US-0157137.
XX PR 03-NOV-1999; 99US-0163280.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX PI WPI; 2001-235357/24.
XX DR P-PSDB; AAG75223.
XX DR
XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX PT useful for preventing, diagnosing and/or treating colorectal cancers -
XX PS Claim 1; Page 3301; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
XX cancer-associated nucleic acid molecules (N) and proteins (P), where
XX the proteins are collectively known as colon cancer antigens. The colon
XX cancer antigens have cytotstatic activity and can be used in gene
XX therapy and vaccine production. N and P may be used in the prevention,
XX diagnosis and treatment of diseases associated with inappropriate P
XX expression. For example, N and P may be used to treat disorders
XX CC associated with decreased expression by rectifying mutations or deletions
XX CC in a patient's genome that affect the activity of P by expressing
XX CC inactive proteins or to supplement the patients own production of P.
XX CC Additionally, N may be used to produce the colon cancer-associated ps,
XX CC by inserting the nucleic acids into a host cell and culturing the cell
XX CC to express the proteins. N and P can be used in the prevention, diagnosis
XX CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
XX CC and AAB77789 represent sequences used in the exemplification of the
XX CC present invention.
XX CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
XX CC missing at time of publication, meaning no sequences are present for
XX CC SEQ ID NO:1027 to 1052, 7921 and 7922.
XX SQ Sequence 895 BP; 327 A; 152 C; 158 G; 255 T; 3 other;

Query Match 2.0%; Score 29; DB 22; Length 895;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 1450 CCAGTAAAAA...AAAAAAAAAAAAAAAA 1478
DB 773 CCAGTAAAAA...AAAAAAAAAAAAAAAA 801

RESULT 22
AAV40521
ID AAV40521 standard; cDNA; 921 BP.
XX
XX AC AAV40521;
XX
XX DT 27-OCT-1998 (first entry)
XX
XX DE Homo sapiens CH27_1 clone secreted protein coding region.
XX
XX KW secreted protein; CH27_1; ds.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT CDS 100..678
XX FT /*tag= a
XX FT /note= "secreted protein"
XX
XX PN WO9830695-A2.
XX PS 16-JUL-1998.

```

```

XX 09-JAN-1998; 98WO-US00543.
XX PF
XX PR 08-JAN-1998; 98US-0004684.
XX PR 09-JAN-1997; 97US-0780814.
XX
XX PA (GEMY ) GENETICS INST INC.
XX
XX AGostino MJ, Jacobs K, Lavallie ER, Mccoy JM, Merberg D;
XX PI Racie LA, Spaulding V, Treacy M;
XX
XX WPI; 1998-413686/35.
XX DR P-PSDB; AAW29660.
XX
XX PT New isolated nucleic acids and secreted proteins - obtained from
XX PT human adult ovary, human foetal kidney, human foetal brain and human
XX PT adult brain cDNA libraries
XX
XX PS Claim 16; Page 68-69; 113pp; English.
XX
XX CC The sequence is that encoding a novel, isolated secreted protein.
XX
XX SQ Sequence 921 BP; 329 A; 162 C; 228 G; 202 T; 0 other;

Query Match 2.0%; Score 29; DB 19; Length 921;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 1450 CCAGTAAAAA...AAAAAAAAAAAAAAAA 1478
DB 818 CCAGTAAAAA...AAAAAAAAAAAAAAAA 846

RESULT 23
AAV38684
ID AAV38684 standard; cDNA; 1018 BP.
XX
XX AC AAV38684;
XX
XX DT 27-OCT-1998 (first entry)
XX
XX DE Homo sapiens SOCS12 cDNA contig 12.2.
XX
XX KW SOCS; suppressor of cytokine signalling; PCR primer;
XX KW autoimmune disease; diagnosis; cancer; treatment;
XX KW cytokine mediated cellular responsiveness; hyperimmunity;
XX KW immunosuppression; allergies; hypertension; contig; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO9820023-A1.
XX
XX PD 14-MAY-1998.
XX
XX PF 31-OCT-1997; 97WO-AU00729.
XX
XX PR 14-FEB-1997; 97AU-0005117.
XX PR 01-NOV-1996; 96AU-0003384.
XX
XX PA (HALL-) HALL INST MEDICAL RES WALTER & ELIZA.
XX
XX PI Alexander WS, Hilton DJ, Metcalf D, Nicholson SE;
XX PI Nicola NA, Richardson RT, Starr R, Viney EM, Willson TA;
XX
XX WPI; 1998-286854/25.
XX
XX PT Suppressor of cytokine signalling proteins - useful to treat
XX PT disease, injury or abnormality involving cytokine mediated cellular
XX PT responsiveness e.g. hyperimmunity, immunosuppression, allergies and
XX PT hypertension
XX
XX PS Claim 14; Page 161-162; 325pp; English.
XX

```

CC The sequence is that of a cDNA contig for a suppressor of cytokine
 CC signalling protein (SOCS). SOCS can be used to screen for naturally
 CC occurring antibodies to SOCS, which may occur, e.g. in some autoimmune
 CC diseases. Alternatively, specific antibodies can be used to
 CC screen for SOCS, which is useful as a knowledge of SOCS levels
 CC may be important for the diagnosis of certain cancers. Soluble
 CC SOCS polypeptides can be used to treat disease, injury or
 CC abnormality involving cytokine mediated cellular responsiveness,
 CC e.g. hyperimmunity, immunosuppression, allergies and hypertension.
 XX

XX Sequence 1018 BP; 306 A; 198 C; 211 G; 303 T; 0 other;

Query Match 2.0%; Score 29; DB 19; Length 1018;

Best Local Similarity 100.0%; Pred. No. 0.025;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1450 CCAGTAAAAA 1478

Db 987 CCAGTAAAAA 1015

RESULT 24

AAV63308 standard; cDNA; 1029 BP.

XX AAV63308;

XX 21-JAN-1999 (first entry)

XX Nucleotide nDiAnk1029 encoding an ankyrin protein.

XX Ankyrin protein; parasitic helminth; heartworm disease;

XX elephantiasis; hydrocele; ss.

XX Dirofilaria immitis.

XX US5824306-A.

XX 20-OCT-1998.

XX 26-FEB-1998; 98US-0031485.

XX 24-APR-1997; 97US-0847429.

XX 26-FEB-1998; 98US-0031485.

XX (HESK-) HESKA CORP.

XX Blehm ES, Tang L;

XX WPI: 1998-593373/50.

XX P-PSDB; AAW70602.

XX Dirofilaria and Brugia ankyrin proteins and antibodies - useful for
 XX protection of animals from disease caused by parasitic helminth

XX Claim 1; Columns 51-54; 84pp; English.

XX The present sequence encodes a Dirofilaria immitis ankyrin protein.

XX The ankyrin protein, or anti-ankyrin antibodies, may be used to

XX protect an animal from disease caused by a parasitic helminth,

XX especially where the disease is heartworm disease, elephantiasis or

XX hydrocele.

XX Sequence 1029 BP; 361 A; 173 C; 249 G; 246 T; 0 other;

Query Match

Best Local Similarity 2.0%; Score 29; DB 19; Length 1029;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1450 CCAGTAAAAA 1478

Db 1001 CCAGTAAAAA 1029

RESULT 25

AAV63000 standard; cDNA; 1029 BP.

XX AAV63000;

XX 15-JAN-1999 (first entry)

XX D. immitis ankyrin nDiAnk1029 cDNA.

XX Ankyrin; helminth; parasite; vaccine; infection;

XX passive immunogen; cytotoxic agent; ss.

XX Dirofilaria immitis.

XX Key Location/Qualifiers

XX CDS 1..814

XX /*tag= a

XX /product= "nDiAnk937"

XX /note= "partial ankyrin sequence, no start codon given"

XX US5827692-A.

XX 27-OCT-1998.

XX 24-APR-1997; 97US-0847429.

XX 24-APR-1997; 97US-0847429.

XX (HESK-) HESKA CORP.

XX Blehm ES, Tang L;

XX WPI: 1998-593992/50.

XX P-PSDB; AAW76770.

XX Nucleic acids encoding ankyrins from helminth parasites - useful for
 XX recombinant production of the proteins for use as vaccines and
 XX treatments against helminth infection

XX Claim 1; Column 51-54; 84pp; English.

XX AAV62996-V63027 encode ankyrin proteins isolated from the helminth
 XX parasites Dirofilaria immitis and Brugia malayi. The nucleic acids and
 XX recombinant products are useful for the recombinant production of the
 XX ankyrin polypeptides. These proteins can then be used as vaccines
 XX against parasitic helminth, e.g. D. immitis or B. malayi. They can also
 XX be used for therapy after infection, and to raise antibodies, also for
 XX use in therapeutics, as passive immunogens, or as therapeutics against
 XX helminths on conjugation to cytotoxic agents. The nucleic acids contained
 XX in viruses, may also be used as viral vaccines, and the nucleic acids
 XX themselves or in vectors may be used as genetic vaccines.

XX Sequence 1029 BP; 361 A; 173 C; 249 G; 246 T; 0 other;

Query Match

Best Local Similarity 2.0%; Score 29; DB 19; Length 1029;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1450 CCAGTAAAAA 1478

Db 1001 CCAGTAAAAA 1029

RESULT 26

AAV63001/C

XX AAV63001 standard; cDNA; 1029 BP.

XX AAV63001;

XX 15-JAN-1999 (first entry)

XX


```
OS  Dirofilaria immitis.
XX  US6063599-A.
XX  16-MAY-2000.
XX  24-APR-1998; 98US-0065474.
XX  24-APR-1997; 97US-0847429.
XX  (HESK-) HESKA CORP.
XX  Blehm ES, Tang L;
PI  WPI: 2000-375493/32.
DR  P-PSDB; RAB11583.
XX  New Dirofilaria and Brugia ankyrin proteins and nucleic acid encoding
PT  them, useful for treating and protecting animals from diseases caused
PT  by parasitic helminths, e.g. heartworm disease, elephantiasis or
PT  hydrocele
XX  Example 1; Column 63-64; 120pp; English.
XX  The invention relates to ankyrin proteins and nucleic acids from the
XX  parasitic helminths Dirofilaria immitis and Brugia malayi. It also
XX  relates to antibodies raised against such ankyrin proteins and to
XX  compounds that inhibit Dirofilaria or Brugia ankyrin function.
XX  Dirofilaria ankyrin cDNAs were isolated from a D. immitis 48 hour
XX  L3 cDNA library using PCR primers based on the sequence of the El
XX  ankyrin from Onchocerca volvulus and the Caenorhabditis elegans ankyrin
XX  UNC-44 genes. Brugia ankyrin cDNAs were isolated from a B. malayi adult
XX  female cDNA library using D. immitis ankyrin and C. elegans UNC-44 PCR
XX  primers. Dirofilaria or Brugia ankyrin proteins and nucleic acids
XX  represent novel targets for anti-helminthic vaccines and drugs. Ankyrin
XX  for protecting animals, particularly dogs, from diseases caused by
XX  nucleic acid molecules, proteins, vaccines and compositions are useful
XX  for protecting animals, particularly dogs, from diseases caused by
XX  parasitic helminths (e.g., heartworm disease, elephantiasis or
XX  hydrocele), as well as for treating the infection. The ankyrin nucleic
XX  acid molecules, proteins, vaccines and compositions of the invention are
XX  especially useful in treating and preventing infections caused by
XX  nematodes (e.g., D. immitis and B. malayi), and ascarid, capillaria,
XX  strongylid, strongyloides, trichostrongyle, and trichurid nematodes and
XX  are also useful against cestodes and trematodes. The therapeutic
XX  compositions may be administered to mammals, including dogs, cats,
XX  humans, ferrets, horses, cattle, sheep, and other pets; economic food
XX  animals; or zoo animals. The ankyrin nucleic acid molecules, proteins and
XX  compounds may also be used as diagnostic reagents to detect infection by
XX  parasitic helminths. Prior art anti-helminthic drugs require repeated
XX  administration, which often leads to the development of resistant
XX  helminth strains that no longer respond to treatment. Such drugs can
XX  also cause harmful side effects in the individual being treated, and a
XX  number of these drugs can only treat the symptoms of a parasitic disease,
XX  being unable to prevent infection by the parasitic helminth. Elucidation
XX  of D. immitis and B. malayi ankyrin protein and DNA sequences facilitates
XX  the development of agents which inhibit ankyrin-mediated parasite
XX  developmental and migratory pathways. Sequences AAA58169-A58196,
XX  AAA58220- AAA58241 and AAA58272-A58277 represent cDNA sequences encoding
XX  D. immitis ankyrin or fragments thereof.
XX  Query Match 2.0%; Score 29; DB 21; Length 1029;
XX  Best Local Similarity 100.0%; Pred. No. 0.025;
XX  Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1450 CCAGTAAACAAAAAAGAAAAAAGAAAAA 1478
DB 29 CCAGTAAACAAAAAAGAAAAAAGAAAAA 1
RESULT 29
AAS72702
```

```
ID  AAS72702 standard; cDNA; 1306 BP.
XX  AAS72702;
XX  13-FEB-2002 (first entry)
XX  DNA encoding novel human diagnostic protein #8506.
XX  Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX  food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX  Homo sapiens.
XX  WO200175067-A2.
XX  11-OCT-2001.
XX  30-MAR-2001; 2001WO-US08631.
XX  31-MAR-2000; 2000US-0540217.
XX  23-AUG-2000; 2000US-0649167.
XX  (HYSE-) HYSEQ INC.
XX  Drmanac RT, Liu C, Tang YT;
XX  WPI: 2001-639362/73.
XX  P-PSDB; ABG08515.
XX  New isolated polynucleotide and encoded polypeptides, useful in
XX  diagnostics, forensics, gene mapping, identification of mutations
XX  responsible for genetic disorders or other traits and to assess
XX  biodiversity
XX  Claim 1; SEQ ID No 8506; 103pp; English.
XX  The invention relates to isolated polynucleotide (I) and
XX  polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX  polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX  and gene mapping, and in recombinant production of (II). The
XX  polynucleotides are also used in diagnostics as expressed sequence tags
XX  for identifying expressed genes. (I) is useful in gene therapy techniques
XX  to restore normal activity of (II) or to treat disease states involving
XX  (II). (II) is useful for generating antibodies against it, detecting or
XX  quantitating a polypeptide in tissue, as molecular weight markers and as
XX  imaging of sites expressing (II). (I) and (II) are useful for treating
XX  disorders involving aberrant protein expression or biological activity.
XX  The polypeptide and polynucleotide sequences have applications in
XX  diagnostics, forensics, gene mapping, identification of mutations
XX  responsible for genetic disorders or other traits to assess biodiversity
XX  and to produce other types of data and products dependent on DNA and
XX  amino acid sequences. AAS64197-AAS94564 represent novel human
XX  diagnostic coding sequences of the invention.
XX  Note: The sequence data for this patent did not appear in the printed
XX  specification, but was obtained in electronic format directly from WIPO
XX  at ftp.wipo.int/pub/published_pct_sequences.
XX  Sequence 1306 BP; 541 A; 235 C; 211 G; 319 T; 0 other;
XX  Query Match 2.0%; Score 29; DB 23; Length 1306;
XX  Best Local Similarity 100.0%; Pred. No. 0.024;
XX  Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1450 CCAGTAAACAAAAAAGAAAAAAGAAAAA 1478
DB 1124 CCAGTAAACAAAAAAGAAAAAAGAAAAA 1152
RESULT 30
AAH17037
ID AAH17037 standard; cDNA; 1902 BP.
XX
```

```

AC RAH17037;
XX
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cDNA sequence SEQ ID NO:16339.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1999; 99JP-0300253.
XX
XX 11-JAN-2000; 2000JP-0118776.
XX
XX 02-MAY-2000; 2000JP-0183767.
XX
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Oka T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the
XX full-length cDNAs
XX
XX Claim 8; SEQ ID 16339; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
XX to the complementary strand of a polynucleotide which comprises one of
XX the 5602 nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in
XX the specification. The primer sets can be used in antisense therapy and
XX in gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
XX represent oligonucleotides, all of which are used in the exemplification
XX of the present invention.
XX
XX Sequence 1902 BP; 605 A; 348 G; 381 G; 568 T; 0 other;
XX
XX Query Match 2.0%; Score 29; DB 22; Length 1902;
XX Best Local Similarity 100.0%; Pred. No. 0.023;
XX Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1450 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
Db 1558 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1586
RESULT 31
AAZ36833
ID AAZ36833 standard; cDNA; 2263 BP.

```

```

XX
XX AAZ36833;
XX
XX 13-MAR-2000 (first entry)
XX
XX cDNA encoding a suppressor of cytokine signalling protein, HSCOP-9.
XX
XX Human; suppressor of cytokine signalling protein; SOCS protein; HSCOP;
XX cancer; leukaemia; lymphoma; diabetes mellitus; Crohn's disease;
XX immune disorder; AIDS; allergy; atherosclerosis; inflammatory disorder;
XX rheumatoid arthritis; irritable bowel syndrome; multiple sclerosis;
XX ulcerative colitis; neurological disorder; Down's syndrome; amnesia;
XX cerebral neoplasm; Huntingdon's disease; viral infection; adenovirus;
XX acute respiratory disease; toga virus; rubella; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 175..1731
XX FT /*tag= a
XX FT /product= "HSCOP-9"
XX
XX WO9961614-A2.
XX
XX 02-DEC-1999.
XX
XX 25-MAY-1999; 99WO-US11497.
XX
XX 28-MAY-1998; 98US-0087104.
XX
XX 17-DEC-1998; 98US-0216006.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Lal P, Hillman JL, Gorgone G, Corley NC, Patterson C, Yue H;
XX Tang YT, Azimzai Y;
XX
XX WPI: 2000-072621/06.
XX P-PSDB; AAY53889.
XX
XX New purified polypeptide encoding human suppressor of cytokine
XX signalling (SOCS) proteins useful for diagnosing, treating or preventing
XX disorders associated with human SOCS proteins
XX
XX Claim 9; Page 89-90; 90pp; English.
XX
XX The present sequence encodes a human suppressor of cytokine signalling
XX (SOCS) protein, designated HSCOP-9. The protein is useful for treating
XX and/or preventing a disorder associated with decreased expression or
XX activity of HSCOP. The protein antagonist is useful for treating and/or
XX preventing a disorder associated with increased expression or activity
XX of HSCOP. The human SOCS proteins and polynucleotides encoding them are
XX useful in the diagnosis, treatment and prevention of cancer such as
XX leukaemia and lymphoma (especially e.g. cancers of the bone, heart and
XX skin), diabetes mellitus, Crohn's disease, immune disorders e.g. AIDS,
XX allergies and atherosclerosis, inflammatory disorders e.g. rheumatoid
XX arthritis, irritable bowel syndrome, multiple sclerosis and ulcerative
XX colitis, neurological disorders e.g. Down's syndrome, amnesia, cerebral
XX neoplasms and Huntingdon's disease and infectious diseases such as
XX those caused by viral infection e.g. adenoviruses (acute respiratory
XX disease) and toga viruses (rubella) as well as those caused by
XX bacterial, fungal, parasitic, protozoal and helminthic infections.
XX
XX Sequence 2263 BP; 650 A; 438 C; 497 G; 678 T; 0 other;
XX
XX Query Match 2.0%; Score 29; DB 21; Length 2263;
XX Best Local Similarity 100.0%; Pred. No. 0.022;
XX Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1450 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
Db 1911 CCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1939

```

```
RESULT 32
ABA09048/c
ID ABA09048 standard; cDNA: 4623 BP.
XX
AC ABA09048;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human ASB-3 protein homologue-encoding cDNA, SEQ ID NO:824.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW hematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; anti-inflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer; ss.
XX
OS Homo sapiens.
XX
PN W0200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX
FA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
WPI; 2001-457740/49.
XX
P-PSDB; ABB11804.
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer -
XX
PS Claim 1; Page 730-731; 1963pp; English.
XX
sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
invention also relates to vectors and recombinant host cells comprising a
nucleotide of the invention, methods of producing the novel polypeptides,
antibodies against the polypeptides, methods of detecting the nucleotides
or polypeptides in a sample, and methods of identifying compounds which
bind to polypeptides of the invention. Although novel, many of the
polypeptides of the invention have homology to known proteins, thereby
giving an insight into their probable biological activities, and hence
potential therapeutic applications. The polypeptides of the invention may
have various activities, including cytokine, cell proliferation or cell
differentiation activities; stem cell growth factor activity;
hematopoiesis regulatory activity; tissue growth activity;
immunomodulatory activity; activin- or inhibin-related activities;
chemotactic or chemokinetic activities; haemostatic, thrombotic or
thrombolytic activities; receptor or ligand activities; or may be
involved in oncogenesis, cancer cell proliferation or metastasis.
Depending on their biological activities, polypeptides and nucleotides of
the invention are useful for preventing, treating or ameliorating medical
conditions, e.g., by protein or gene therapy. Such conditions include
cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
proliferative retinopathy, atherosclerosis, coronary heart disease,
arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
vascular growth. Polypeptides involved with tissue regeneration and
repair (or nucleic acids encoding them) may be used to promote wound
healing (e.g., of burns, incisions and ulcers), while those with
immunomodulatory activities may be used in the treatment of viral,
bacterial and fungal infections in addition to immune disorders.
Polypeptides with growth factor activity may be used in cell cultures to
promote cell growth. For example, such polypeptides may be used to
manipulate stem cells in culture to give rise to neuroepithelial cells
that can be used to augment or replace cells damaged by illness,
autoimmune disease or accidental damage. The polypeptides and nucleotides
may also be used in the diagnosis of the above conditions, and in drug
screening techniques. The present sequence represents a cDNA encoding a
novel human polypeptide of the invention.
XX
SQ Sequence 4623 BP; 1371 A; 971 C; 978 G; 1303 T; 0 other;
Query Match 2.0%; Score 29; DB 22; Length 4623;
Best Local Similarity 100.0%; Pred. No. 0.02;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1450 CCAGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 1478
|
DB 183 CCAGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 155
|
RESULT 33
AAV63314
ID AAV63314 standard; cDNA: 5503 BP.
XX
AC AAV63314;
XX
DT 21-JAN-1999 (first entry)
XX
DE Full length nucleotide nDAnk5503 encoding an ankyrin protein.
KW Ankyrin protein; parasitic helminth; heartworm disease;
KW elephantiasis; hydrocele; ss.
XX
OS Dirofilaria immitis.
XX
FH Key Location/Qualifiers
FT CDS 51..5288
FT /tag= a
FT /product= ankyrin protein
XX
US5824306-A.
XX
PD 20-OCT-1998.
XX
PF 26-FEB-1998; 98US-0031485.
XX
PR 24-APR-1997; 97US-0847429.
PR 26-FEB-1998; 98US-0031485.
XX
FA (HESK-) HESKA CORP.
XX
PI Blehm ES, Tang L;
XX
WPI; 1998-593373/50.
XX
DR P-PSDB; AAW70608.
XX
PT Dirofilaria and Brugia ankyrin proteins and antibodies - useful for
PT protection of animals from disease caused by parasitic helminth
XX
PS Claim 1; Columns 95-108; 84pp; English.
XX
The present sequence encodes a full length Dirofilaria immitis ankyrin
protein. The ankyrin protein, or anti-ankyrin antibodies, may be used to
protect an animal from disease caused by a parasitic helminth,
especially where the disease is heartworm disease, elephantiasis or
hydrocele.
XX
SQ Sequence 5503 BP; 1662 A; 1152 C; 1259 G; 1429 T; 1 other;
```



```

Query Match      2.0%; Score 29; DB 19; Length 5503;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1450 CCAGTAAAAA...AAAAAAAAAAAAA 1478
    |||||||...|||||
DB 5475 CCAGTAAAAA...AAAAAAAAAAAAA 5503

RESULT 34
AAV63020
ID AAV63020 standard; cDNA: 5503 BP.
XX
AC AAV63020;
XX
XX 15-JAN-1999 (first entry)
XX
DE D. immitis ankyrin cDNA complementary to nDiAnk5503.
XX
XX Ankyrin; helminth; parasite; vaccine; therapy; infection;
KW passive immunogen; cytotoxic agent; ss.
XX
OS Dirofilaria immitis.
XX
PN US827692-A.
XX
XX 27-OCT-1998.
XX
XX 24-APR-1997; 97US-0847429.
XX
XX 24-APR-1997; 97US-0847429.
XX
XX (HESK-) HESKA CORP.
XX
XX Blehm ES, Tang L;
XX
XX WPI; 1998-593992/50.
XX
XX Nucleic acids encoding ankyrins from helminth parasites - useful for
PT recombinant production of the proteins for use as vaccines and
PT treatments against helminth infection
XX
XX Claim 1; Column 117-124; 84pp; English.
XX
XX AAV62996-V63027 encode ankyrin proteins isolated from the helminth
CC parasites Dirofilaria immitis and Brugia malayi. The nucleic acids and
CC recombinant products are useful for the recombinant production of the
CC ankyrin polypeptides. These proteins can then be used as vaccines
CC against parasitic helminth, e.g. D. immitis or B. malayi. They can also
CC be used for therapy after infection, and to raise antibodies, also for
CC use in therapeutics, as passive immunogens, or as therapeutics against
CC helminths on conjugation to cytotoxic agents. The nucleic acids contained
CC in viruses, may also be used as viral vaccines, and the nucleic acids
CC themselves or in vectors may be used as genetic vaccines.
XX
XX Sequence 5503 BP; 1429 A; 1259 C; 1152 G; 1429 T; 1 other;
SQ

Query Match      2.0%; Score 29; DB 19; Length 5503;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1450 CCAGTAAAAA...AAAAAAAAAAAAA 1478
    |||||||...|||||
DB 29 CCAGTAAAAA...AAAAAAAAAAAAA 1

RESULT 36
AAV58193
ID AAV58193 standard; cDNA: 5503 BP.
XX
AC AAV58193;
XX
XX 23-OCT-2000 (first entry)
XX
XX D. immitis ankyrin cDNA coding strand, nDiAnk5503.
DE
XX
XX Ankyrin; parasitic helminth; filarid nematode; heartworm disease;
KW elephantiasis; hydrocele; vaccine; antibody; antihelminthic; ss.
XX
XX Dirofilaria immitis.
XX
XX US6063599-A.
XX

```

PD 16-MAY-2000.
 XX 24-APR-1998; 98US-0065474.
 XX 24-APR-1997; 97US-0847429.
 XX (HESK-) HESKA CORP.
 XX Blehm ES, Tang L;
 XX WPI; 2000-375493/32.
 DR P-PSDB; AAB11589.
 XX New Dirofilaria and Brugia ankyrin proteins and nucleic acid encoding
 PT them, useful for treating and protecting animals from diseases caused
 PT by parasitic helminths, e.g. heartworm disease, elephantiasis or
 PT hydrocele
 XX
 PS Example 1: Column 103-118; 120pp; English.
 XX
 CC The invention relates to ankyrin proteins and nucleic acids from the
 CC parasitic helminths Dirofilaria immitis and Brugia malayi. It also
 CC relates to antibodies raised against such ankyrin proteins and to
 CC compounds that inhibit Dirofilaria or Brugia ankyrin function.
 CC Dirofilaria ankyrin cDNAs were isolated from a D. immitis 48 hour
 CC L3 cDNA library using PCR primers based on the sequence of the El
 CC ankyrin from Onchocerca volvulus and the Caenorhabditis elegans ankyrin
 CC UNC-44 genes. Brugia ankyrin cDNAs were isolated from a B. malayi adult
 CC female cDNA library using D. immitis ankyrin and C. elegans UNC-44 PCR
 CC primers. Dirofilaria or Brugia ankyrin proteins and nucleic acids
 CC represent novel targets for anti-helminthic vaccines and drugs. Ankyrin
 CC nucleic acid molecules, proteins, vaccines and compositions are useful
 CC for protecting animals, particularly dogs, from diseases caused by
 CC parasitic helminths (e.g., heartworm disease, elephantiasis or
 CC hydrocele), as well as for treating the infection. The ankyrin nucleic
 CC acid molecules, proteins, vaccines and compositions of the invention are
 CC especially useful in treating and preventing infections caused by filarial
 CC nematodes (e.g., D. immitis and B. malayi), and ascarid, capillaria,
 CC strongylid, strongyloides, trichostrongyle, or trichurid nematodes and
 CC are also useful against cestodes and trematodes. The therapeutic
 CC compositions may be administered to mammals, including dogs, cats,
 CC humans, ferrets, horses, cattle, sheep, and other pets; economic food
 CC animals; or zoo animals. The ankyrin nucleic acid molecules, proteins and
 CC compounds may also be used as diagnostic reagents to detect infection by
 CC parasitic helminths. Prior art anti-helminthic drugs require repeated
 CC administration, which often leads to the development of resistant
 CC helminth strains that no longer respond to treatment. Such drugs can
 CC also cause harmful side effects in the individual being treated, and a
 CC number of these drugs can only treat the symptoms of a parasitic disease,
 CC being unable to prevent infection by the parasitic helminth. Elucidation
 CC of D. immitis and B. malayi ankyrin protein and DNA sequences facilitates
 CC the development of agents which inhibit ankyrin-mediated parasite
 CC development and migratory pathways. Sequences AAA58169-A58196,
 CC AAA58220-AAA58241 and AAA58272-A58277 represent cDNA sequences encoding
 CC D. immitis ankyrin or fragments thereof.
 XX
 SQ Sequence 5503 BP; 1662 A; 1152 C; 1259 G; 1429 T; 1 other;
 Query Match 2.0%; Score 29; DB 21; Length 5503;
 Best Local Similarity 100.0%; Pred. No. 0.019;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1450 CCAGTAAAAAATAAAAAAAAAAAAAA 1478
 ||||||||||||||||||||||||||||
 Db 5475 CCAGTAAAAAATAAAAAAAAAAAAAA 5503
 RESULT 37
 AAA58194/c
 ID AAA58194 standard; cDNA; 5503 BP.
 XX
 AC AAA58194;
 XX

DT 23-OCT-2000 (first entry)
 XX D. immitis ankyrin nDiAnk503 complementary strand.
 DE
 XX Ankyrin; parasitic helminth; filariid nematode; heartworm disease;
 KW elephantiasis; hydrocele; vaccine; antibody; antihelminthic; ss.
 XX
 XX Dirofilaria immitis.
 OS
 XX US6063599-A.
 PN
 XX 16-MAY-2000.
 PD
 XX 24-APR-1998; 98US-0065474.
 PF
 XX 24-APR-1997; 97US-0847429.
 PR
 XX (HESK-) HESKA CORP.
 PA
 XX Blehm ES, Tang L;
 PI
 XX WPI; 2000-375493/32.
 DR P-PSDB; AAB11589.
 DR
 XX New Dirofilaria and Brugia ankyrin proteins and nucleic acid encoding
 PT them, useful for treating and protecting animals from diseases caused
 PT by parasitic helminths, e.g. heartworm disease, elephantiasis or
 PT hydrocele
 XX
 PS Disclosure; Column 127-134; 120pp; English.
 XX
 CC The invention relates to ankyrin proteins and nucleic acids from the
 CC parasitic helminths Dirofilaria immitis and Brugia malayi. It also
 CC relates to antibodies raised against such ankyrin proteins and to
 CC compounds that inhibit Dirofilaria or Brugia ankyrin function.
 CC Dirofilaria ankyrin cDNAs were isolated from a D. immitis 48 hour
 CC L3 cDNA library using PCR primers based on the sequence of the El
 CC ankyrin from Onchocerca volvulus and the Caenorhabditis elegans ankyrin
 CC UNC-44 genes. Brugia ankyrin cDNAs were isolated from a B. malayi adult
 CC female cDNA library using D. immitis ankyrin and C. elegans UNC-44 PCR
 CC primers. Dirofilaria or Brugia ankyrin proteins and nucleic acids
 CC represent novel targets for anti-helminthic vaccines and drugs. Ankyrin
 CC nucleic acid molecules, proteins, vaccines and compositions are useful
 CC for protecting animals, particularly dogs, from diseases caused by
 CC parasitic helminths (e.g., heartworm disease, elephantiasis or
 CC hydrocele), as well as for treating the infection. The ankyrin nucleic
 CC acid molecules, proteins, vaccines and compositions of the invention are
 CC especially useful in treating and preventing infections caused by filarial
 CC nematodes (e.g., D. immitis and B. malayi), and ascarid, capillaria,
 CC strongylid, strongyloides, trichostrongyle, or trichurid nematodes and
 CC are also useful against cestodes and trematodes. The therapeutic
 CC compositions may be administered to mammals, including dogs, cats,
 CC humans, ferrets, horses, cattle, sheep, and other pets; economic food
 CC animals; or zoo animals. The ankyrin nucleic acid molecules, proteins and
 CC compounds may also be used as diagnostic reagents to detect infection by
 CC parasitic helminths. Prior art anti-helminthic drugs require repeated
 CC administration, which often leads to the development of resistant
 CC helminth strains that no longer respond to treatment. Such drugs can
 CC also cause harmful side effects in the individual being treated, and a
 CC number of these drugs can only treat the symptoms of a parasitic disease,
 CC being unable to prevent infection by the parasitic helminth. Elucidation
 CC of D. immitis and B. malayi ankyrin protein and DNA sequences facilitates
 CC the development of agents which inhibit ankyrin-mediated parasite
 CC development and migratory pathways. Sequences AAA58169-A58196,
 CC AAA58220-AAA58241 and AAA58272-A58277 represent cDNA sequences encoding
 CC D. immitis ankyrin or fragments thereof.
 XX
 SQ Sequence 5503 BP; 1429 A; 1259 C; 1152 G; 1662 T; 1 other;
 Query Match 2.0%; Score 29; DB 21; Length 5503;
 Best Local Similarity 100.0%; Pred. No. 0.019;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAAAA 1478
 DB 29 CCAGTAAAAA 1

RESULT 38

AAT72978
 ID AAT72978 standard; DNA; 54 BP.
 XX AC AAT72978;
 XX DT 27-FEB-1998 (first entry)
 XX DE Knot probe p306 used to detect amplification of M. tuberculosis DNA.
 XX KW DnaJ gene; knot probe; nucleic acid detection; oligonucleotide probe;
 XX KW termolecular complex; chemiluminescence signal; homogenous format; ss.
 XX OS Synthetic.
 XX OS Mycobacterium tuberculosis.
 XX FH Key Location/Qualifiers
 XX FT modified_base 1
 XX FT /*tag= a
 XX FT /*note= '5' A labelled with digoxin"
 XX FT modified_base 54
 XX FT /*tag= b
 XX FT /*note= '3' G labelled with phosphate"
 XX PN WO9723647-A1.
 XX PD 03-JUL-1997.
 XX PF 20-DEC-1996; 96WO-US19751.
 XX PR 22-DEC-1995; 95US-0009090.
 XX PA (BEHW) BEHRINGERWERKE AG.
 XX PI (ULLM/) ULLMAN EF.
 XX PI Kurn N, Lin C, Liu YP, Patel RD, Rose SJ, Ullman EF;
 XX DR WPI; 1997-351082/32.
 XX PT Amplification and detection of target polynucleotide in homogenous
 XX PT format - minimises the number and complexity of steps and reagents
 XX PS Example 2C; Page 51; 82pp; English.

XX Knot probes AAT72977-78 were used to detect the amplification of
 CC M. tuberculosis genomic DNA (see AAT72976). The present sequence
 CC represents knot probe p306. These oligonucleotides were used to
 CC demonstrate a novel method for amplifying and detecting a target
 CC polynucleotide sequence. This method comprises providing in combination
 CC a medium suspected of containing the target sequence, all amplification
 CC reagents and two oligonucleotide probes capable of binding to a single
 CC strand of the product of the amplification product. At least one of the
 CC probes has two sequences which are non-contiguous and can bind to
 CC contiguous or non-contiguous sites on the single strand or can bind to
 CC non-contiguous sites on the single strand. Both of the probes hybridise
 CC to one of the strands to form a termolecular complex, which is then
 CC detected (e.g. in the present example, using chemiluminescence signals).
 CC The method is used to amplify and detect nucleic acids in a homogenous
 CC format. The method minimises the number and complexity of steps and
 CC reagents.

XX Sequence 54 BP; 33 A; 6 C; 8 G; 7 T; 0 other;

Query Match 1.9%; Score 28; DB 18; Length 54;
 Best Local Similarity 100.0%; Pred. No. 0.096;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1451 CAGTAAAAA 1478

DB 12 CAGTAAAAA 39

RESULT 39

AAAX16888/C
 ID AAAX16888 standard; DNA; 63 BP.
 XX AC AAAX16888;
 XX DT 27-APR-1999 (first entry)
 XX DE Rat cytochrome P450 gene PCR primer #10.
 XX KW Internal standard; quantitative analysis; rat; cytochrome P-450; primer;
 XX KW amplification; RT-PCR; detection; drug; gene expression; liver; ss.
 XX OS Synthetic.
 XX OS Rattus sp.
 XX PN JPL1018800-A.
 XX PD 26-JAN-1999.
 XX PF 04-JUL-1997; 97JP-0193402.
 XX PR 04-JUL-1997; 97JP-0193402.
 XX PA (TAKI) TAKARA SHUZO CO LTD.
 XX DR WPI; 1999-161091/14.
 XX PT Internal standard nucleic acid - useful for determination of the
 XX PT molecular species of rat cytochrome P-450
 XX PS Example 1; Page 12; 13pp; Japanese.

XX Primers AAAX16879-X16888 are used to PCR amplify part of the rat
 CC cytochrome P-450 (Cyp) gene. The invention relates to a cyp RNA sequence
 CC (AAAX16860) useful as an internal standard in the quantitative analysis of
 CC rat cytochrome P-450 mRNA molecular species by nucleic acid
 CC amplification. The standard gives an amplified product of a chain length
 CC distinguishable from the amplified product derived from the mRNA of the
 CC molecular species to be analysed. The size of the products amplified by
 CC primers AAAX16879-X16888 are compared to the standards amplified by
 CC primers AAAX16861-X16876. Detection methods using the internal standard
 CC can easily distinguish and determine the effect of a drug on the
 CC expression of P-450 molecular species in a rat liver.

XX Sequence 63 BP; 7 A; 7 C; 12 G; 37 T; 0 other;
 Query Match 1.9%; Score 28; DB 20; Length 63;
 Best Local Similarity 100.0%; Pred. No. 0.094;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1451 CAGTAAAAA 1478
 DB 46 CAGTAAAAA 19

RESULT 40

AAAS60388/C
 ID AAAS60388 standard; cDNA; 161 BP.
 XX AC AAAS60388;
 XX DT 29-JAN-2002 (first entry)

XX Human cancer agent-resistance marker #262.

XX Human: cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;
 KW squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;
 KW lymphocytic leukaemia; lymphoma; plasmocytoma; reticulum cell sarcoma;

KW Hodgkin's disease; glioma; ss.
 XX Homo sapiens.
 XX WO200179556-A2.
 XX PD 25-OCT-2001.
 XX PF 13-APR-2001; 2001WO-US12132.
 XX PR 14-APR-2000; 2000US-197538P.
 XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PI Lillie J, Brown JL, Bolt A, Van Huffel C;
 XX WPI; 2001-602933/68.
 DR Novel nucleic acid, used as a marker to determine the effectiveness of
 PT using TAXOL to treat cancer cell growth in individuals.
 XX Claim 1; Page 205; 527pp; English.
 XX The invention relates to 1046 novel nucleic acids which are used as
 CC markers for determining the sensitivity of a cancer cell to the
 CC anticancer agent TAXOL. Cancer cells can be treated with TAXOL when
 CC they are shown to express one of the 242 sensitivity markers or the
 CC cells are shown not to express one of the 804 resistance markers.
 CC The methods can be used to determine the effectiveness of TAXOL
 CC in the treatment of cancer cell growth in an individual. The markers
 CC can be used as targets in developing anti-cancer agents such as
 CC chemotherapeutic compounds. The markers can also be used as targets in
 CC developing treatments for cancer, particularly those cancers which
 CC display resistance to agents and exhibit expression of the markers. The
 CC anticancer agents developed by the novel method can be used to treat
 CC cancer. Probes based on the markers can be used to detect transcripts or
 CC genomic sequences corresponding to the markers, in the identification of
 CC cells or tissues which mis-express the protein. Cancers which may
 CC be targeted include carcinoma (e.g. squamous cell carcinoma),
 CC sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic leukaemia),
 CC lymphoma, plasmacytoma, reticulum cell sarcoma, Hodgkin's disease and
 CC tumours (e.g. glioma). The present sequence is one of the 1046
 CC novel cancer cell markers.
 XX
 SQ Sequence 161 BP; 31 A; 22 C; 29 G; 72 T; 7 other;
 Query Match 1.9%; Score 28; DB 22; Length 161;
 Best Local Similarity 100.0%; Pred. No. 0.082;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1451 CAGTAAAAA...AAAAAAAAAAAAAAAA 1478
 DB 73 CAGTAAAAA...AAAAAAAAAAAAAAAA 46
 RESULT 41
 AAS07705/c
 ID AAS07705 standard; DNA; 322 BP.
 XX
 AC AAS07705;
 XX
 DT 23-OCT-2001 (first entry)
 XX
 DE Cervical cancer pre-malignant condition DNA marker #4.
 XX
 KW Cervical cancer; pre-malignant condition marker; tumorigenesis; CIN; ds;
 KW Cervical intraepithelial neoplasia; squamous intraepithelial lesions;
 KW polymerase chain reaction; PCR; enzyme linked immunosorbent assay; ELISA;
 KW anticancer therapy; carcinogen; antisense inhibition.
 XX
 OS Homo sapiens.
 XX
 PN WO200142792-A2.

XX 14-JUN-2001.
 XX PD 08-DEC-2000; 2000WO-US33311.
 XX PF 08-DEC-1999; 99US-0169811.
 XX PR 21-DEC-1999; 99US-0171330.
 XX PR 14-MAR-2000; 2000US-0189113.
 XX PR 31-MAR-2000; 2000US-0193943.
 XX PR 12-MAY-2000; 2000US-0203772.
 XX PR 09-JUN-2000; 2000US-0210820.
 XX PR 21-JUL-2000; 2000US-0220113.
 XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PI Schlegel R, Deeds J, Berger A, Zhao X;
 XX WPI; 2001-367889/38.
 DR Cervical cancer protein markers useful for the diagnosis, prevention
 PT and treatment of cervical cancers, especially cervical intraepithelial
 PT neoplasia or squamous intraepithelial lesions -
 XX
 XX 20; Page 405; 436pp; English.
 XX The sequence represents a cervical cancer pre-malignant condition marker.
 CC These markers encode proteins that are over expressed during
 CC tumorigenesis. The proteins and their corresponding nucleic acid
 CC sequences can therefore be used for the diagnosis, prevention and
 CC treatment of cervical cancers, particularly cervical intraepithelial
 CC neoplasia (CIN) and squamous intraepithelial lesions (SIL). The sequences
 CC may be used as markers in diagnostic assays to detect cancerous
 CC conditions (e.g. by polymerase chain reaction (PCR) or enzyme linked
 CC immunosorbent assay (ELISA)), to monitor the efficacy of anticancer
 CC therapies and to identify anticancer or carcinogenic compounds. The level
 CC of expression of a marker in a patient sample is compared with the normal
 CC level of expression of the marker in a control non-cervical sample is
 CC sample, whereby a significant difference indicates that the patient is
 CC afflicted with cervical cancer or a pre-malignant condition. The
 CC expression of these proteins may be inhibited by antisense inhibition for
 CC the treatment of cancers. They may be used in this way for the treatment
 CC of cervical intraepithelial neoplasia or squamous intraepithelial
 CC lesions.
 XX
 SQ Sequence 322 BP; 93 A; 72 C; 56 G; 98 T; 3 other;
 Query Match 1.9%; Score 28; DB 22; Length 322;
 Best Local Similarity 100.0%; Pred. No. 0.074;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1451 CAGTAAAAA...AAAAAAAAAAAAAAAA 1478
 DB 55 CAGTAAAAA...AAAAAAAAAAAAAAAA 28
 RESULT 42
 AAL18635/c
 ID AAL18635 standard; cDNA; 330 BP.
 XX
 AC AAL18635;
 XX
 DT 07-DEC-2001 (first entry)
 XX
 DE Human breast cancer expressed polynucleotide 11092.
 XX
 KW Human; breast cancer; cell marker; cytostatic; ss.
 OS Homo sapiens.
 XX
 PN WO200151628-A2.
 XX
 PD 19-JUL-2001.
 XX

```

PF 10-JAN-2001; 2001WO-US00798.
XX
PR 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer
XX
XX Claim 1; Page 1976; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
SQ Sequence 330 BP; 101 A; 57 C; 91 G; 81 T; 0 other;
XX
Query Match 1.9%; Score 28; DB 22; Length 330;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1451 CAGTAAAAAAGAAAAAAGAAAAAAGAAAAA 1478
DB 35 CAGTAAAAAAGAAAAAAGAAAAAAGAAAAA 8
XX
RESULT 43
AAL16827
ID AAL16827 standard; cDNA; 333 BP.
XX
AC AAL16827;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 9284.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US00798.
XX
PR 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX Lillie J, Xu Y, Wang Y, Steinmann K;
PI
XX

```

```

DR WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer
XX
XX Claim 1; Page 1668-1669; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
SQ Sequence 333 BP; 78 A; 90 C; 58 G; 104 T; 3 other;
XX
Query Match 1.9%; Score 28; DB 22; Length 333;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1451 CAGTAAAAAAGAAAAAAGAAAAAAGAAAAA 1478
DB 299 CAGTAAAAAAGAAAAAAGAAAAAAGAAAAA 326
XX
RESULT 44
AAL26029
ID AAL26029 standard; cDNA; 349 BP.
XX
AC AAL26029;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 18486.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US00798.
XX
PR 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
PI
XX
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer
XX
XX Claim 1; Page 3411; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX

```

```
CC activity.
XX
SQ Sequence 349 BP; 78 A; 97 C; 65 G; 106 T; 3 other;

Query Match      1.9%; Score 28; DB 22; Length 349;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1451 CAGTAAAAA...AAAAAAAAA 1478
      |||...|||
Db 319 CAGTAAAAA...AAAAAAAAA 346

RESULT 45
AAQ60162/C
ID AAQ60162 standard; DNA; 382 BP.
XX
AC AAQ60162;
XX
DT 16-MAR-1994 (first entry)
XX
DE Human brain Expressed Sequence Tag EST02150.
XX
KW Gene transcription product; genetic markers; tagging; in vivo;
KW transcription; mapping; locations; chromosomes; chromosomal; ss.
XX
CS Homo sapiens.
XX
PN W09316178-A.
XX
PD 19-AUG-1993.
XX
PF 12-FEB-1993; 93WO-US01294.
XX
PR 12-FEB-1992; 92US-0837195.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Adams MD, Moreno RF, Venter CJ;
XX
DR WPI; 1993-272882/34.
XX
PT Enriched oligonucleotides and corresp. sequences - used as
PT markers for human genes transcribed in-vivo, facilitate tagging
PT of most human genes
XX
PS Example 4; Page 291; 500pp; English.
XX
CC The Expressed Sequence Tag was isolated from a human brain cDNA
CC library as part of a large set of ESTs which can be used as markers
CC for human genes transcribed in vivo. They can be used to facilitate
CC tagging of most human genes, for mapping locations of expressed genes
CC on chromosomes, for individual or forensic identification, for mapping
CC locations of disease-associated genes, for identification of tissue
CC type, and for prepn. of antisense sequences, probes and constructs.
CC EST02150 has a "poor" coding probability as evaluated using the
CC coding-region prediction program CRM. See also AAQ59041-Q61440.
XX
SQ Sequence 382 BP; 94 A; 66 C; 90 G; 129 T; 3 other;

Query Match      1.9%; Score 28; DB 14; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.073;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1451 CAGTAAAAA...AAAAAAAAA 1478
      |||...|||
Db 39 CAGTAAAAA...AAAAAAAAA 12

Search completed: November 5, 2002, 13:52:00
Job time : 279.457 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 11:01:37 ; Search time 36.249 Seconds
(without alignments)
10015.359 Million cell updates/sec

```
Title: US-09-805-311-7
Perfect score: 1478
Sequence: 1 CQACCCACGCTCCGQAAA.....AAAAAAAAAAAAAAAAAAAA 1478
```

Scoring table: OLIGO_NUC
Gapop 50.0 , Gapext 50.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

```
Database :
1: /cgn2.6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2.6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2.6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2.6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2.6/ptodata/1/ina/PTGUS_COMB.seq.*
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description	
		Match	%					
1	1478	100.0	1478	4	US-09-426-557-7		Sequence 7, Appli	
2	894	60.5	1463	4	US-09-426-557-1		Sequence 1, Appli	
3	850	57.5	1381	4	US-09-426-557-5		Sequence 5, Appli	
4	842	57.0	1541	4	US-09-426-557-3		Sequence 3, Appli	
5	29	2.0	1029	1	US-09-031-485-6		Sequence 6, Appli	
c	29	2.0	1029	1	US-09-031-485-8		Sequence 8, Appli	
6	29	2.0	1029	1	US-08-847-429A-6		Sequence 6, Appli	
c	29	2.0	1029	1	US-08-847-429A-8		Sequence 8, Appli	
8	29	2.0	1029	3	US-09-065-474-6		Sequence 6, Appli	
9	29	2.0	1029	3	US-09-065-474-8		Sequence 8, Appli	
c	10	29	2.0	1029	3	US-09-031-485-32		Sequence 32, Appl
11	29	2.0	5503	1	US-09-031-485-34		Sequence 34, Appl	
c	12	29	2.0	5503	1	US-08-847-429A-32		Sequence 32, Appl
13	29	2.0	5503	1	US-08-847-429A-34		Sequence 34, Appl	
c	14	29	2.0	5503	3	US-09-065-474-32		Sequence 32, Appl
15	29	2.0	5503	3	US-09-065-474-34		Sequence 34, Appl	
c	16	29	2.0	5503	3	US-09-065-474-34		Sequence 34, Appl
17	28	1.9	54	2	US-08-771-624B-24		Sequence 24, Appl	
18	28	1.9	1602	4	US-09-333-423-3		Sequence 3, Appli	
19	28	1.9	1633	4	US-09-119-788-1		Sequence 1, Appli	
20	28	1.9	4137	3	US-09-221-235-1		Sequence 1, Appli	
21	28	1.9	4137	3	US-09-221-928-1		Sequence 1, Appli	
22	28	1.9	4137	3	US-09-221-527-1		Sequence 1, Appli	
23	28	1.9	4137	3	US-09-221-236-1		Sequence 1, Appli	
24	28	1.9	4137	3	US-09-221-416-1		Sequence 1, Appli	
25	28	1.9	4137	3	US-09-221-245-1		Sequence 1, Appli	
26	28	1.9	4137	4	US-09-163-115-1		Sequence 1, Appli	
27	28	1.9	4137	4	US-09-221-528-1		Sequence 1, Appli	

ALIGNMENTS

```

RESULT 1
US-09-426-557-7
; Sequence 7, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramo B.
; TITLE OF INVENTION: Maize Rad2/FEN1
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09-
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112-
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows vers
; SEQ ID NO 7
; LENGTH: 1478
; TYPE: DNA
; ORGANISM: zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (97)...(1233)
US-09-426-557-7

```

Query Match	100.0%;	Score 1478;	DB 4;	Length 1478;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1478;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
Qy	1	CGACCCACGGTCCGGGAATAGTCTCGGGTTCGGGGTTTCTTGGCCACATCCCGGCTCAGC	60	
Db	1	CGACCCACGCGTCCGGGAATAGTCTCGGGTTCGGGGTTTCTTGGCCACATCCCGGCTCAGC	60	
Qy	61	CGCGCGCCGCGACCCGCCACAGCCCGCAGACGAGATGGGCATCAAGGTTTTCAGCGAAA	120	
Db	61	CGCGCGCCGCGACCCGCCACACCCCGCAGACGAGATGGGCATCAAGGTTTTCAGCGAAA	120	
Qy	121	CTGCTGGCGGACAAATGCGCCCAAGCGGATGAAGGAGCAGAAGTTCGAGAGCTACTTCGGC	180	
Db	121	CTGCTGGCGGACAAATGCGCCCAAGCGGATGAAGGAGCAGAAGTTCGAGAGCTACTTCGGC	180	
Qy	181	CGCAAAATCGCCGTCCGACCCAGCATGAGCATCTACCAAGTTCCCTGATGATGTTGGAAGG	240	
Db	181	CGCAAAATCGCCGTCCGACCCAGCATGAGCATCTACCAAGTTCCCTGATGATGTTGGAAGG	240	
Qy	241	ACAGGCATGGAACCTCTCAAAATGAAGCTGGTGAAGTCAGTACTATTTGCAAGGAATG	300	
Db	241	ACAGGCATGGAACCTCTCAAAATGAAGCTGGTGAAGTCAGTACTATTTGCAAGGAATG	300	
Qy	301	TTCAACCGGACAAATAGATTACTTGAAGCGGGAATCAAGCCAGTTTATGTTTGTGATGGC	360	

```
Db 301 TTCAACCGGACAAATAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTTTGATGCC 360
Qy 361 AAGCCTCTGTATGATGAAGAAACAAAGAACTTGTCTAAAGATACTCAAAAGAGATGATGCA 420
Db 361 AAGCCTCTGTATGATGAAGAAACAAAGAACTTGTCTAAAGATACTCAAAAGAGATGATGCA 420
Qy 421 ACCAAGATCTGACTGAGGCAGCTAGAGTAGGAGATAAGATGCGATTGAAAATTGAGC 480
Db 421 ACCAAGATCTGACTGAGGCAGCTAGAGTAGGAGATAAGATGCGATTGAAAATTGAGC 480
Qy 481 AAGAGACTGTAAAGGTCAAGAGCAACACACACAGAGATGTTAAACGACTATTAAAGACTT 540
Db 481 AAGAGACTGTAAAGGTCAAGAGCAACACACACAGAGATGTTAAACGACTATTAAAGACTT 540
Qy 541 ATGGGGTTCCTGTTCTAGAGCCACTTCTGAAGCAGAGCAGAAATGTGAGCCCTTTGC 600
Db 541 ATGGGGTTCCTGTTCTAGAGCCACTTCTGAAGCAGAGCAGAAATGTGAGCCCTTTGC 600
Qy 601 ATAAAGATAAGGTGTTCCTGTTCTGCTCAGAGATAAGGACTCCCTACTTTTGGGCT 660
Db 601 ATAAAGATAAGGTGTTCCTGTTCTGCTCAGAGATAAGGACTCCCTACTTTTGGGCT 660
Qy 661 CCAGGTTCCTGCTCATTTAATGATCCAAAGTTCCAAGAAATACCTGTGATGGAATTT 720
Db 661 CCAGGTTCCTGCTCATTTAATGATCCAAAGTTCCAAGAAATACCTGTGATGGAATTT 720
Qy 721 GATGTTGCCAAGGTTTGGAGAGCTTGAATCACCATTGAGACAGTTTCATTTGTGC 780
Db 721 GATGTTGCCAAGGTTTGGAGAGCTTGAATCACCATTGAGACAGTTTCATTTGTGC 780
Qy 781 ATCTGTGTGGATGTGACTATTATGATAGCATCAAGGATATCGGGGCAACAGCTG 840
Db 781 ATCTGTGTGGATGTGACTATTATGATAGCATCAAGGATATCGGGGCAACAGCTG 840
Qy 841 AAATTTATTCGTCATCATGGTCCATAGAAAGCATCTTGGAGATCTTTAATAAGACAGA 900
Db 841 AAATTTATTCGTCATCATGGTCCATAGAAAGCATCTTGGAGATCTTTAATAAGACAGA 900
Qy 901 TATCAAAATTCCTGAGACTGGGCTTACCAAGAAAGCTCGACGCTTGTTCAGAGAGCCTAAT 960
Db 901 TATCAAAATTCCTGAGACTGGGCTTACCAAGAAAGCTCGACGCTTGTTCAGAGAGCCTAAT 960
Qy 961 GTCACTTTGATATTCTGAGCTAAATGGACTGCACCTGATGAGGAGGTTCTCATAGT 1020
Db 961 GTCACTTTGATATTCTGAGCTAAATGGACTGCACCTGATGAGGAGGTTCTCATAGT 1020
Qy 1021 TTCTGTGTAAGATAATGTTTCAATGAAGATCGGGTGACAAAGGCGCATAGAGAATC 1080
Db 1021 TTCTGTGTAAGATAATGTTTCAATGAAGATCGGGTGACAAAGGCGCATAGAGAATC 1080
Qy 1081 AAATCTGCCAAGATAAATCGTCGCAAGGAAGACTCGAGTCTTTTTCAGCCCAACTGCC 1140
Db 1081 AAATCTGCCAAGATAAATCGTCGCAAGGAAGACTCGAGTCTTTTTCAGCCCAACTGCC 1140
Qy 1141 ACCACATCAGCAGCTTAAACGGAAGGAGACTTCGGATAAACAAAGCAAGCGAGCTGCG 1200
Db 1141 ACCACATCAGCAGCTTAAACGGAAGGAGACTTCGGATAAACAAAGCAAGCGAGCTGCG 1200
Qy 1201 AACAAAGAAACAAGGCTGGTGGAAAGAAAGAAATATCTTGATGCTTGATACACTA 1260
Db 1201 AACAAAGAAACAAGGCTGGTGGAAAGAAAGAAATATCTTGATGCTTGATACACTA 1260
Qy 1261 CGACTACGAAGCAGCGGTGGCATGATCATTCTGCGCTAGATATTAACTCCCTGTTTTA 1320
Db 1261 CGACTACGAAGCAGCGGTGGCATGATCATTCTGCGCTAGATATTAACTCCCTGTTTTA 1320
Qy 1321 ACTCAGACCTTTGGTGAAGTTTGGCCATGTTTCAAGCTGGGGTAAGTTAGTTGTTG 1380
Db 1321 ACTCAGACCTTTGGTGAAGTTTGGCCATGTTTCAAGCTGGGGTAAGTTAGTTGTTG 1380
Qy 1381 AAGAGATTGCTGTACCAAGTAACAAACTTATCGCTGTTTTTTTACTCTTCTCCTTTCAA 1440
Db 1381 AAGAGATTGCTGTACCAAGTAACAAACTTATCGCTGTTTTTTTACTCTTCTCCTTTCAA 1440
```

```
Qy 1441 GTATGTATGCCAGTAAAAAAGAAAAAAGAAAAAAGAAAAA 1478
Db 1441 GTATGTATGCCAGTAAAAAAGAAAAAAGAAAAAAGAAAAA 1478

RESULT 2
US-09-426-557-1
; Sequence 1, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1463
; TYPE: DNA
; ORGNISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (85)...(1221)
US-09-426-557-1
```

Query Match 60.5%; Score 894; DB 4; Length 1463;

Best Local Similarity 99.4%; Pred. No. 0;
Matches 1244; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy 32 CGCGGTTTCTTGGCCACTCCGGCTCAGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCAG 91
Db 20 CGCGGTTTCTTGGCCACTCCGGCTCAGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCAG 79
Qy 92 ACGAGATGGGCATCAAGGGTTTGACGAAACTGCTGCGGAGCAATGCGGCCCAAGCGCATGA 151
Db 80 ACGAGATGGGCATCAAGGGTTTGACGAAACTGCTGCGGAGCAATGCGGCCCAAGCGCATGA 139
Qy 152 AGGAGCAGAGTTTCGAGAGCTACTTCGGCCGCAAAATCGCGTCGACGCCACAGCATGAGCA 211
Db 140 AGGAGCAGAGTTTCGAGAGCTACTTCGGCCGCAAAATCGCGTCGACGCCACAGCATGAGCA 199
Qy 212 TCTACCAAGTTCCCTGATAGTAGTTTGAAGGACAGGCGATGGAACCTCTCACAATGAAGCTG 271
Db 200 TATACCAAGTTCCCTGATAGTAGTTTGAAGGACAGGCGATGGAACCTCTCACAATGAAGCTG 259
Qy 272 GTGAAGTCACCTAGTCATTTGCAAGGAATGTTCAACGGGACAAATAGATTACTTGAAGCGG 331
Db 260 GTGAAGTCACCTAGTCATTTGCAAGGAATGTTCAACGGGACAAATAGATTACTTGAAGCGG 319
Qy 332 GAATCAAGCCAGTTTATGTTTGTGATGGCAAGCCCTCCTGATATGAAGAAACAAGAACTTG 391
Db 320 GAATCAAGCCAGTTTATGTTTGTGATGGCAAGCCCTCCTGATATGAAGAAACAAGAGCTTG 379
Qy 392 CTAAAAATGACTCAAAAGAGATGATGCAACCAAGAGATCTGACTGAGCGCATAGAGTAG 451
Db 380 CTAAAAATGACTCAAAAGAGATGATGCAACCAAGAGATCTGACTGAGCGCATAGAGTAG 439
Qy 452 GAGATAAAGATGCGATTGAAAAATTCAGCAAGAGGAGCTGTAAAGCTCAACAGGCAACACA 511
Db 440 GAGATAAAGATGCGATTGAAAAATTCAGCAAGAGGAGCTGTAAAGCTCAACAGGCAACACA 499
Qy 512 ACGAAGATTGTAACGACTATTAAAGACTTATGGGGTTCCTGTTGTAGAGCACCTTCTG 571
Db 500 ACGAAGATTGTAACGCGCTATTAAAGACTTATGGGGTTCCTGTTGTAGAGCACCTTCTG 559
Qy 572 AAGCAGAGCAGAAATGTGAGCCCTTTTGCATAAAGCATAAAGGTGTTTCGCTGTTGCTTCA 631
Db 560 AAGCAGAGCAGAAATGTGAGCCCTTTTGCATAAAGCATAAAGGTGTTTCGCTGTTGCTTCA 619
```



```
QY 632 AAGATGAAGACATCCCTTACTTTTGGGGCTCCACGGTCCCTTCGTCAATTTAATGGATCCAA 691
Db 620 AAGATATGAGACTCCCTTACTTTTGGGGCTCCACGGTCCCTTCGTCAATTTAATGGATCCAA 679
QY 692 GTTCCAGAAATACCTGTGATGGAATTTGATGTTGGCAAGGTTTTCGAGGAGCTTGAAC 751
Db 680 GTTCCAGAAATACCTGTGATGGAATTTGATGTTGGCAAGGTTTTCGAGGAGCTTGAAC 739
QY 752 TCACCATGAGACAGTTCAATGATTTGTCATCCCTGTGTGGATGACTATTTGATGATAGA 811
Db 740 TCACCATGAGACAGTTCAATGATTTGTCATCCCTGTGTGGATGACTATTTGATGATAGA 799
QY 812 TCAAAAGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCACATGGGTCCATAGAAA 871
Db 800 TCAAAAGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCACATGGGTCCATAGAAA 859
QY 872 GCATCTGGAGATCTTAATAAGACAGATATCAAAATTCCTGGAGACTGCCTTACCAAG 931
Db 860 GCATCTGGAGATCTTAATAAGACAGATATCAAAATTCCTGGAGACTGCCTTACCAAG 919
QY 932 AAGCTCGACGCTTGTTCAGGAGCCTTAATGTCACCTTTGGATATTCCTGAGCTAAATGGA 991
Db 920 AAGCTCGACGCTTGTTCAGGAGCCTTAATGTCACCTTTGGATATTCCTGAGCTAAATGGA 979
QY 992 CTGCACCTGATGAGGAGGTCTCATAGTTTCTGGTAAAGATAATGTTTCAATGAAG 1051
Db 980 CTGCACCTGATGAGGAGGTCTCATAGTTTCTGGTAAAGATAATGTTTCAACGAAG 1039
QY 1052 ATCGGGTGACAAGGCCATAGAGAGATCAAAATCTGCCAAGATAAATCTGCCAAGGA 1111
Db 1040 ATCGGGTGACAAGGCCATAGAGAGATCAAAATCTGCCAAGATAAATCTGCCAAGGA 1099
QY 1112 GACTCGAGTCTTTTCAAGCCACTGCCACACATCAGCACCGCTAAAACGGAAGSAGA 1171
Db 1100 GACTCGAGTCTTTTCAAGCCACTGCCACACATCAGCACCGCTAAAACGGAAGSAGA 1159
QY 1172 CTTTCGATAAACAGCAGGCGCTGCGAACAAGAAACAAAGGCTGTGGAAAGAGA 1231
Db 1160 CTTTCGATAAACAGCAGGCGCTGCGAACAAGAAACAAAGGCTGTGGAAAGAGA 1219
QY 1232 AATAATCTTGGATGCTGATGATACAACTACGACTACGAAAGCAGCGGTGSC 1282
Db 1220 AATAATCTTGGATGCTGATGATACAACTACGACTACGAAAGCAGCGGTGSC 1270
```

RESULT 3

```
US-09-426-557-5
; Sequence 5, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37)...(1173)
US-09-426-557-5
```

Query Match 57.5%; Score 850; DB 4; Length 1381;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1200; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY 76 GCCACAGCCGCCACGACGAGATGGGCATCAAGGGTTTGACGAAACTGCTGGCGGACAAT 135
Db 16 GCCACAGCCGCCACGACGAGATGGGCATCAAGGGTTTGACGAAACTGCTGGCGGACAAT 75
QY 136 GCGCCCAAGGCGATGAAGGAGCAGAGTTCGAGAGCTACTTTCGGCCGCAAAATCCGCGTC 195
Db 76 GCGCCCAAGGCGATGAAGGAGCAGAGTTCGAGAGCTACTTTCGGCCGCAAAATCCGCGTC 135
QY 196 GAGCCAGCATGAGCATCTACCAGTTTCCCTGATAGTAGTTGGAAGGACAGGCGATGAAACT 255
Db 136 GAGCCAGCATGAGCATATACCAGTTTCCCTGATAGTAGTTGGAAGGACAGGCGATGAAACT 195
QY 256 CTCAAAATCAAGCTGGTGAAGTCACTACTAGTCAATTTGCAAGGAATGTTCAACGGGACAATA 315
Db 196 CTCAAAATCAAGCTGGTGAAGTCACTACTAGTCAATTTGCAAGGAATGTTCAACGGGACAATA 255
QY 316 AGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGGATGGCAAGCCTCCCTGATATG 375
Db 256 AGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGGATGGCAAGCCTCCCTGATATG 315
QY 376 AAGAACCAAGACTTCTTAAAGATCTCAAAAGAGATGATGCAACCAAGAGATCTGACT 435
Db 316 AAGAACCAAGACTTCTTAAAGATCTCAAAAGAGATGATGCAACCAAGAGATCTGACT 375
QY 436 GAGSCAGTAGAGGTAGGATGAAGATGCGATTGAAAATTTGAGCAAGGAGACTGTAAG 495
Db 376 GAGSCAGTAGAGGTAGGATGAAGATGCGATTGAAAATTTGAGCAAGGAGACTGTAAG 435
QY 496 GTCAAGAGGCAACACAAAGCAATGTTAAACGACTATTTAAGACTTATGGGGGTTCCCTGTT 555
Db 436 GTCAAGAGGCAACACAAAGCAATGTTAAACGACTATTTAAGACTTATGGGGGTTCCCTGTT 495
QY 556 GTAGAGCAGCTTCTGAAGCAGAGCAATGTGAGCGCTTTGTCATAAAGCAATAAGGTG 615
Db 496 GTAGAGCAGCTTCTGAAGCAGAGCAATGTGAGCGCTTTGTCATAAAGCAATAAGGTG 555
QY 616 TTGCGTGTTCCTCAGAAATGAAGACTCCCTTACTTTTGGGGCTCCAGGTTCCCTTCGT 675
Db 556 TTGCGTGTTCCTCAGAAATGAAGACTCCCTTACTTTTGGGGCTCCAGGTTCCCTTCGT 615
QY 676 CATTTAATGATCCAAAGTTCCAAAGAAATACCTGTGATGGAAATTTGATGTTGCCAAGGTT 735
Db 616 CATTTAATGATCCAAAGTTCCAAAGAAATACCTGTGATGGAAATTTGATGTTGCCAAGGTT 675
QY 736 TTGGAGGAGCTTGAACATCCATGACAGTTCATGATTTGTGATCCTGTGATGATGT 795
Db 676 TTGGAGGAGCTTGAACATCCATGACAGTTCATGATTTGTGATCCTGTGATGATGT 735
QY 796 GACTATTGTGATAGCATCAAGGTATCGGGGGCAACAGCTCTGAAACTTTATTCGTCAA 855
Db 736 GACTATTGTGATAGCATCAAGGTATCGGGGGCAACAGCTCTGAAACTTTATTCGTCAA 795
QY 856 CATGGGTCCATAGAAAGCATCTTTGGAGATCTTAATAAAGACAGATATCAAAATTCCTGAG 915
Db 796 CATGGGTCCATAGAAAGCATCTTTGGAGATCTTAATAAAGACAGATATCAAAATTCCTGAG 855
QY 916 GACTGGCTTTACCAAGAGCTGACGCTTTGTCAGAGGAGCTTAATGTCACTTCGATATT 975
Db 856 GACTGGCTTTACCAAGAGCTGACGCTTTGTCAGAGGAGCTTAATGTCACTTCGATATT 915
QY 976 CTTGAGCTAAATGGAGCTGACCTGATGAGGAGGTCTCATAAGTTTCCCTGGTAAAGAT 1035
Db 916 CTTGAGCTAAATGGAGCTGACCTGATGAGGAGGTCTCATAAGTTTCCCTGGTAAAGAT 975
QY 1036 AATGGTTTCAATGAAGTGGGTGACAAAGCCATAGAGAAGATCAAAATCTGCCAAGAAAT 1095
Db 976 AATGGTTTCAACGAAGATCGGGTGACAAAGCCATAGAGAAGATCAAAATCTGCCAAGAAAT 1035
QY 1096 AATCTGTCGAAGGAAGACTTCGAGTCCCTTTTCAAGGCAACTGCCACACATCAGCACCG 1155
Db 1036 AATCTGTCGAAGGAAGACTTCGAGTCCCTTTTCAAGGCAACTGCCACACATCAGCACCG 1095
```

QY 1156 CTAACGGAAGGAGACTTCGGATAAACAAGCAGCTGCGAACAAGAAACAAAG 1215
|||||
Db 1096 CTAACGGAAGGAGACTTCGGATAAACAAGCAGCTGCGAACAAGAAACAAAG 1155
|||||
QY 1216 CTTGGTGAAGAAATAATCTTGGATCTGTATGTACACTACGACTAGCAAGCAG 1275
|||||
Db 1156 CTTGGTGAAGAAATAATCTTGGATCTGTATGTACACTACGACTAGCAAGCAG 1215
|||||
QY 1276 CGGTGGC 1282
|||||
Db 1216 CGGTGGC 1222
|||||
RESULT 4
US-09-426-557-3
; Sequence 3, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)....(1215)
US-09-426-557-3
Query Match 57.0%; Score 842; DB 4; Length 1541;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1242; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 33 GCGGTTTCTTGGCCCACTCGGCTCAGCGCGCCGCCGCCGCCGCCGCCGCCGCGCAGA 92
|||||
Db 15 GCGGTTTCTTGGCCCACTCGGCTCAGCGCGCCGCCGCCGCCGCCGCCGCCGCGCAGA 74
|||||
QY 93 CGAGATGGGCATCAAGGGTTTGACAAACTGCTGGCGGACAAATGCGCCCAAGCGCATGAA 152
|||||
Db 75 CGAGATGGGCATCAAGGGTTTGACAAACTGCTGGCGGACAAATGCGCCCAAGCGCATGAA 134
|||||
QY 153 GGAGCAGAAGTTCGAGAGCTACTTCGGCCGCAAAATCGCGCGCAGCAGCATGAGCAT 212
|||||
Db 135 GGAGCAGAAGTTCGAGAGCTACTTCGGCCGCAAAATCGCGCGCAGCAGCATGAGCAT 194
|||||
QY 213 CTACCAAGTTCCTGATAGTACTTGGAAAGGACAGCATGGAACCTCTCAAAATGAAGCTGG 272
|||||
Db 195 ATACCAAGTTCCTGATAGTACTTGGAAAGGACAGCATGGAACCTCTCAAAATGAAGCTGG 254
|||||
QY 273 TGAAGTCACATAGTCATTTCGAAGGAATGTTCAACCGGACAAATAAGATTACTTGAAGCGGG 332
|||||
Db 255 TGAAGTCACATAGTCATTTCGAAGGAATGTTCAACCGGACAAATAAGATTACTTGAAGCGGG 314
|||||
QY 333 AATCAAGCCAGTTTATGTTTTGATGGCAAGCTCCTGATATGAAGAACAAGACTTGC 392
|||||
Db 315 AATCAAGCCAGTTTATGTTTTGATGGCAAGCTCCTGATATGAAGAACAAGACTTGC 374
|||||
QY 393 TAAAGATACCTAAAGAGATGATCAACCAAGATCTGACTGAGGAGTAGAGTAGG 452
|||||
Db 375 TAAAGATACCTAAAGAGATGATCAACCAAGATCTGACTGAGGAGTAGAGTAGG 434
|||||
QY 453 AGATAAAGATCGGATTGAAAAATTTGAGCAAGAGGACTGTAAAGGTCAAGGCAACACAA 512
|||||
Db 435 AGATAAAGATCGGATTGAAAAATTTGAGCAAGAGGACTGTAAAGGTCAAGGCAACACAA 494
|||||

QY 513 CGAAGATTGTAACGACTATTAAAGACTTTATGGGGTTTCTGTGTAGAGCACCCTTCTGA 572
|||||
Db 495 CGAAGATTGTAACGACTATTAAAGACTTTATGGGGTTTCTGTGTAGAGCACCCTTCTGA 554
|||||
QY 573 AGCAGAACGAGAAATGTGAGCCCTTTGTCATAAAGATAGGTGTCTGCTGTGCTTCTGA 632
|||||
Db 555 AGCAGAACGAGAAATGTGAGCCCTTTGTCATAAAGATAGGTGTCTGCTGTGCTTCTGA 614
|||||
QY 633 AGATAAGGACTCCCTTACTTTTGGGGCTCCACGGTTCCCTCGTCAATTAATGGATCCAAG 692
|||||
Db 615 AGATAAGGACTCCCTTACTTTTGGGGCTCCACGGTTCCCTCGTCAATTAATGGATCCAAG 674
|||||
QY 693 TTCCAAGAAATACCTGTGATGGAATTTGATGTTGCCAAGGTTTTGGAGGAGCTTGAAC 752
|||||
Db 675 TTCCAAGAAATACCTGTGATGGAATTTGATGTTGCCAAGGTTTTGGAGGAGCTTGAAC 734
|||||
QY 753 CACCATGGACCAAGTTCAATTTGTCATCTGTGTGATGTGACTATTGTGATAGCAT 812
|||||
Db 735 CACCATGGACCAAGTTCAATTTGTCATCTGTGTGATGTGACTATTGTGATAGCAT 794
|||||
QY 813 CAAAGTATCGGGGGGCAACAGCTCTGAAACTTATTCGTCAACATGGGTCCTAGAAAG 872
|||||
Db 795 CAAAGTATCGGGGGGCAACAGCTCTGAAACTTATTCGTCAACATGGGTCCTAGAAAG 854
|||||
QY 873 CATCTTGGAGAACTTAATAAAGACAGATATCAAAATTCCTGAGGACTGGCCTTACCAAGA 932
|||||
Db 855 CATCTTGGAGAACTTAATAAAGACAGATATCAAAATTCCTGAGGACTGGCCTTACCAAGA 914
|||||
QY 933 AGCTGACCTTGTTCAGAGGAGCTTAATTCACATTTGGATATTCCTGAGCTAAAATGGAC 992
|||||
Db 915 AGCTGACCTTGTTCAGAGGAGCTTAATTCACATTTGGATATTCCTGAGCTAAAATGGAC 974
|||||
QY 993 TGCACCTGATGAGGAGGCTCTATAAGTTTCTGTGTAAGATTAATGTTTCAATGAAGA 1052
|||||
Db 975 TGCACCTGATGAGGAGGCTCTATAAGTTTCTGTGTAAGATTAATGTTTCAACGAAGA 1034
|||||
QY 1053 TCGGGTGCAAGGCCATAGAGAATCAAACTCTGCCAAGAAATAATCTGTCGCAAGGAAG 1112
|||||
Db 1035 TCGGGTGCAAGGCCATAGAGAATCAAACTCTGCCAAGAAATAATCTGTCGCAAGGAAG 1094
|||||
QY 1113 ACTCGAGTCTTTTCAAGCCAACTGCCACCACTACAGCAGCTGCTGCAAGGAGGAGAC 1172
|||||
Db 1095 ACTCGAGTCTTTTCAAGCCAACTGCCACCACTACAGCAGCTGCTGCAAGGAGGAGAC 1154
|||||
QY 1173 TCGGATAAACAAGGAGGAGCTGCGAACAAGAAACAAGGCTGCTGGAAGAAGAA 1232
|||||
Db 1155 TCGGATAAACAAGGAGGAGCTGCGAACAAGAAACAAGGCTGCTGGAAGAAGAA 1214
|||||
QY 1233 ATAACTTGGATGCTGTGATGTAACACTACGAAAGCAGCGGTGGC 1282
|||||
Db 1215 ATAACTTGGATGCTGTGATGTAACACTACGAAAGCAGCGGTGGC 1264
|||||

RESULT 5

US-09-031-485-6
; Sequence 6, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: PROTEIN, AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEIN, AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEIN, AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEIN, AND BRUGIA ANKYRIN
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Hesk Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031.485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847.429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1029 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..811
; US-09-031-485-6
```

```
Query Match 2.0%; Score 29; DB 1; Length 1029;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1450 CCAGTAAAAA 1478
Db 1001 CCAGTAAAAA 1029
```

```
RESULT 6
US-09-031-485-8/C
; Sequence 8, Application US/09031485
; Patent No. 5824306
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031.485
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/847.429
; FILING DATE: 24-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
```

```
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1029 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-031-485-8
; Query Match 2.0%; Score 29; DB 1; Length 1029;
; Best Local Similarity 100.0%; Pred. No. 0.0016;
; Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; QY 1450 CCAGTAAAAA 1478
; Db 29 CCAGTAAAAA 1
; RESULT 7
; US-08-847-429A-6
; Sequence 6, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847.429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1029 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..811
; US-08-847-429A-6
; Query Match 2.0%; Score 29; DB 1; Length 1029;
; Best Local Similarity 100.0%; Pred. No. 0.0016;
; Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1450 CCAGTAAAAA 1478
 Db 1001 CCAGTAAAAA 1029

RESULT 8

US-08-847-429A-8/C
 ; Sequence 8, Application US/08847429A
 ; Patent No. 5827692
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Liang
 ; APPLICANT: Blehm, E. Scot
 ; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
 ; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
 ; TITLE OF INVENTION: USES THEREOF
 ; NUMBER OF SEQUENCES: 85
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Carol Talkington Verser, Ph.D.
 ; ADDRESSEE: Heska Corporation
 ; STREET: 1825 Sharp Point Drive
 ; CITY: Fort Collins
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80525
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: WordPerfect for Windows, Version 7.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/847,429A
 ; FILING DATE: 24-APR-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Verser, Carol Talkington
 ; REGISTRATION NUMBER: 37,459
 ; REFERENCE/DOCKET NUMBER: HW-5
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 970/493-7272
 ; TELEFAX: 970/484-9505
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1029 nucleotides
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-847-429A-8

Query Match 2.0%; Score 29; DB 1; Length 1029;
 Best Local Similarity 100.0%; Pred. No. 0.0016;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1450 CCAGTAAAAA 1478
 Db 29 CCAGTAAAAA 1

RESULT 9

US-09-065-474-6
 ; Sequence 6, Application US/09065474
 ; Patent No. 6063599
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Liang
 ; APPLICANT: Blehm, E. Scot
 ; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
 ; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
 ; TITLE OF INVENTION: USES THEREOF
 ; NUMBER OF SEQUENCES: 171
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Carol Talkington Verser, Ph.D.
 ; ADDRESSEE: Heska Corporation
 ; STREET: 1825 Sharp Point Drive

; CITY: Fort Collins
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80525
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: WordPerfect for Windows, Version 7.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/065,474
 ; FILING DATE: 24-APR-1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Verser, Carol Talkington
 ; REGISTRATION NUMBER: 37,459
 ; REFERENCE/DOCKET NUMBER: HW-5-C1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 970/493-7272
 ; TELEFAX: 970/484-9505
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1029 nucleotides
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 2..811
 ; US-09-065-474-6

Query Match 2.0%; Score 29; DB 3; Length 1029;
 Best Local Similarity 100.0%; Pred. No. 0.0016;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1450 CCAGTAAAAA 1478
 Db 1001 CCAGTAAAAA 1029

RESULT 10

US-09-065-474-8/C
 ; Sequence 8, Application US/09065474
 ; Patent No. 6063599
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Liang
 ; APPLICANT: Blehm, E. Scot
 ; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
 ; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
 ; TITLE OF INVENTION: USES THEREOF
 ; NUMBER OF SEQUENCES: 171
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Carol Talkington Verser, Ph.D.
 ; ADDRESSEE: Heska Corporation
 ; STREET: 1825 Sharp Point Drive
 ; CITY: Fort Collins
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80525
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: WordPerfect for Windows, Version 7.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/065,474
 ; FILING DATE: 24-APR-1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Verser, Carol Talkington
 ; REGISTRATION NUMBER: 37,459
 ; REFERENCE/DOCKET NUMBER: HW-5-C1

TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1029 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-065-474-8

Query Match 2.0%; Score 29; DB 3; Length 1029;

Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAAAA 1478
Db 29 CCAGTAAAAA 1

RESULT 11

US-09-031-485-32
Sequence 32, Application US/09031485
Patent No. 5824306
GENERAL INFORMATION:

APPLICANT: Tang, Liang
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:

ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031.485
FILING DATE:

CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/847.429
FILING DATE: 24-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505

INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 5503 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 51..5285
US-09-031-485-32

Query Match 2.0%; Score 29; DB 1; Length 5503;
Best Local Similarity 100.0%; Pred. No. 0.0014;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1450 CCAGTAAAAA 1478
Db 5475 CCAGTAAAAA 5503

RESULT 12

US-09-031-485-34/c
Sequence 34, Application US/09031485
Patent No. 5824306
GENERAL INFORMATION:

APPLICANT: Tang, Liang
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:

ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/031.485
FILING DATE:

CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/847.429
FILING DATE: 24-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5

TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 5503 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-031-485-34

Query Match 2.0%; Score 29; DB 1; Length 5503;
Best Local Similarity 100.0%; Pred. No. 0.0014;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1450 CCAGTAAAAA 1478
Db 29 CCAGTAAAAA 1

RESULT 13

US-08-847-429A-32
Sequence 32, Application US/08847429A
Patent No. 5827692
GENERAL INFORMATION:

APPLICANT: Tang, Liang
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
TITLE OF INVENTION: USES THEREOF

```
;
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5503 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 51...5285
; US-08-847-429A-32
;
; Query Match 2.0%; Score 29; DB 1; Length 5503;
; Best Local Similarity 100.0%; Pred. No. 0.0014;
; Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1450 CCAGTAAAAA...AAAAAAAAA 1478
; |||||
; Db 5475 CCAGTAAAAA...AAAAAAAAA 5503
;
; RESULT 14
; US-08-847-429A-34/c
; Sequence 34, Application US/08847429A
; Patent No. 5827692
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,429A
; FILING DATE: 24-APR-1997
```

```
;
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5503 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-847-429A-34
;
; Query Match 2.0%; Score 29; DB 1; Length 5503;
; Best Local Similarity 100.0%; Pred. No. 0.0014;
; Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1450 CCAGTAAAAA...AAAAAAAAA 1478
; |||||
; Db 29 CCAGTAAAAA...AAAAAAAAA 1
;
; RESULT 15
; US-09-065-474-32
; Sequence 32, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang
; APPLICANT: Blehm, E. Scot
; TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 171
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/065,474
; FILING DATE: 24-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-5-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5503 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 51...5285
; US-09-065-474-32
```

Query Match 2.0%; Score 29; DB 3; Length 5503;
 Best Local Similarity 100.0%; Pred. No. 0.0014;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1450 CCAGTAAAAA 1478
 Db 5475 CCAGTAAAAA 5503

RESULT 16

US-09-065-474-34/c
 ; Sequence 34, Application US/09065474
 ; Patent No. 6063599
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Liang
 ; APPLICANT: Blehm, E. Scott
 ; TITLE OF INVENTION: DIOFILARIA AND BRUGIA ANKYRIN
 ; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
 ; TITLE OF INVENTION: US5 THEREOF
 ; NUMBER OF SEQUENCES: 171
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Carol Talkington Verser, Ph.D.
 ; ADDRESSEE: Heska Corporation
 ; STREET: 1825 Sharp Point Drive
 ; CITY: Fort Collins
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80525
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: WordPerfect for Windows, Version 7.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/065,474
 ; FILING DATE: 24-APR-1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Verser, Carol Talkington
 ; REGISTRATION NUMBER: 37,459
 ; REFERENCE/DOCKET NUMBER: HW-5-C1
 ; TELEPHONE: 970/493-7272
 ; TELEFAX: 970/484-9505
 ; INFORMATION FOR SEQ ID NO: 34:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5503 nucleotides
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-09-065-474-34

Query Match 2.0%; Score 29; DB 3; Length 5503;
 Best Local Similarity 100.0%; Pred. No. 0.0014;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1450 CCAGTAAAAA 1478
 Db 29 CCAGTAAAAA 1

RESULT 17

US-08-771-624B-24
 ; Sequence 24, Application US/08771624B
 ; Patent No. 5914230
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Yen ping
 ; APPLICANT: Patel, Rajesh D.
 ; APPLICANT: Kurn, Nuriith
 ; APPLICANT: Lin, Claire
 ; APPLICANT: Rose, Samuel J.
 ; APPLICANT: Ullman, Edwin F.

; TITLE OF INVENTION: Homogeneous Amplification and Detection
 ; TITLE OF INVENTION: Of Nucleic Acids
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Behring Diagnostics GmbH, c/o Dade Behring Inc.
 ; STREET: 1717 Deerfield Road
 ; CITY: Deerfield
 ; STATE: Illinois
 ; COUNTRY: US
 ; ZIP: 60015-0778

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/771,624B
 ; FILING DATE: 20-DEC-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/009,090
 ; FILING DATE: 22-DEC-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Ruzsala, Lois K.
 ; REGISTRATION NUMBER: 39,074
 ; REFERENCE/DOCKET NUMBER: 1030
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (847) 267-5364
 ; TELEFAX: (847) 267-6024
 ; INFORMATION FOR SEQ ID NO: 24:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 54 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FRAGMENT TYPE: C-terminal
 ; US-08-771-624B-24

Query Match 1.9%; Score 28; DB 2; Length 54;
 Best Local Similarity 100.0%; Pred. No. 0.0055;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA 1478
 Db 12 CAGTAAAAA 39

RESULT 18

US-09-333-423-3
 ; Sequence 3, Application US/09333423
 ; Patent No. 6265636
 ; GENERAL INFORMATION:
 ; APPLICANT: Randall, Douglas
 ; APPLICANT: Thelen, Jay
 ; APPLICANT: Miernyk, Jan
 ; APPLICANT: Muszynski, Michael
 ; APPLICANT: Sewalc, Vincent
 ; TITLE OF INVENTION: Pyruvate Dehydrogenase Kinase
 ; FILE REFERENCE: 0818
 ; CURRENT APPLICATION NUMBER: US/09/333,423
 ; CURRENT FILING DATE: 1999-06-15
 ; EARLIER APPLICATION NUMBER: 60/089,998
 ; EARLIER FILING DATE: 1998-06-19
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 1602
 ; TYPE: DNA

```
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)...(1169)
US-09-333-423-3

Query Match          1.9%; Score 28; DB 4; Length 1602;
Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA...AAAAAAAAA 1478
Db 1572 CAGTAAAAA...AAAAAAAAA 1599

RESULT 19
US-09-119-788-1
; Sequence 1, Application US/09119788
; Patent No. 6166193
; GENERAL INFORMATION:
; APPLICANT: Yanagisawa, Masashi
; TITLE OF INVENTION: CDNA CLONE MY1 THAT ENCODES
;       A NOVEL HUMAN 7-TRANSMEMBRANE RECEPTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: United States of America
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,788
; FILING DATE: 21-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/053,790
; FILING DATE: 25-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T
; REGISTRATION NUMBER: 30,954
; REFERENCE/DOCKET NUMBER: GH50029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5515
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-09-119-788-1

Query Match          1.9%; Score 28; DB 4; Length 1633;
Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA...AAAAAAAAA 1478
Db 1602 CAGTAAAAA...AAAAAAAAA 1629

RESULT 20
US-09-221-235-1
; Sequence 1, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
US-09-221-235-1

Query Match          1.9%; Score 28; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA...AAAAAAAAA 1478
Db 4104 CAGTAAAAA...AAAAAAAAA 4131

RESULT 21
US-09-221-928-1
; Sequence 1, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: 09/221,928
; CURRENT FILING DATE: 1998-12-28
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
US-09-221-928-1

Query Match          1.9%; Score 28; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA...AAAAAAAAA 1478
Db 4104 CAGTAAAAA...AAAAAAAAA 4131

RESULT 22
US-09-221-527-1
; Sequence 1, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: 09/221,527
```



```
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-527-1

Query Match
Best Local Similarity 100.0%; Pred. No. 0.0037; Length 4137;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA 1478
Db 4104 CAGTAAAAA 4131

RESULT 23
US-09-221-236-1
; Sequence 1, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-236-1

Query Match
Best Local Similarity 100.0%; Pred. No. 0.0037; Length 4137;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA 1478
Db 4104 CAGTAAAAA 4131

RESULT 24
US-09-221-416-1
; Sequence 1, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-416-1

Query Match
Best Local Similarity 100.0%; Pred. No. 0.0037; Length 4137;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA 1478
Db 4104 CAGTAAAAA 4131

RESULT 25
US-09-221-245-1
; Sequence 1, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-245-1

Query Match
Best Local Similarity 100.0%; Pred. No. 0.0037; Length 4137;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA 1478
Db 4104 CAGTAAAAA 4131

RESULT 26
US-09-163-115-1
; Sequence 1, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-163-115-1

Query Match
Best Local Similarity 100.0%; Pred. No. 0.0037; Length 4137;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA 1478
Db 4104 CAGTAAAAA 4131

RESULT 27
US-09-163-115-1
; Sequence 1, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-163-115-1

Query Match
Best Local Similarity 100.0%; Pred. No. 0.0037; Length 4137;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA 1478
Db 4104 CAGTAAAAA 4131
```

; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-163-115-1

Query Match 1.9%; Score 28; DB 4; Length 4137;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA 1478
|
Db 4104 CAGTAAAAA 4131

RESULT 27

US-09-221-528-1
; Sequence 1, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-528-1

Query Match 1.9%; Score 28; DB 4; Length 4137;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA 1478
|
Db 4104 CAGTAAAAA 4131

RESULT 28

US-09-593-553-1
; Sequence 1, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-593-553-1

Query Match 1.9%; Score 28; DB 4; Length 4137;
Best Local Similarity 100.0%; Pred. No. 0.0037;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1451 CAGTAAAAA 1478
|
Db 4104 CAGTAAAAA 4131

RESULT 29

US-09-221-237-1
; Sequence 1, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,237
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-237-1

Query Match 1.9%; Score 28; DB 4; Length 4137;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1451 CAGTAAAAA 1478
|
Db 4104 CAGTAAAAA 4131

RESULT 30

US-09-338-907-198
; Sequence 198, Application US/09338907
; Patent No. 6265546
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bouqueleret, Lydie
; TITLE OF INVENTION: PROSTATE CANCER GENE
; FILE REFERENCE: GENSET.18CPICP
; CURRENT APPLICATION NUMBER: US/09/338,907
; CURRENT FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: 08/496,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; EARLIER APPLICATION NUMBER: 09/218,207
; EARLIER FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 198
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1..47
; OTHER INFORMATION: polymorphic fragment 4-4-187
; NAME/KEY: allele
; LOCATION: 24

```
; OTHER INFORMATION: polymorphic base A
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1..23
; OTHER INFORMATION: potential microsequencing oligo 4-4-187.misl
; NAME/KEY: primer_bind
; LOCATION: 25..47
; OTHER INFORMATION: complement potential microsequencing oligo 4-4-187.mis2
US-09-338-907-198

Query Match          1.8%; Score 27; DB 4; Length 47;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1452 AGTAAAAA 1478
Db 21 AGTAAAAA 47

RESULT 31
US-09-218-207-198
; Sequence 198, Application US/09218207
; Patent No. 6346381
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Ilya, Chumakov
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: prostate cancer gene
; FILE REFERENCE: GENSET.018CP1
; CURRENT APPLICATION NUMBER: US/09/218,207
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 08/996,306
; EARLIER FILING DATE: 1997-12-22
; EARLIER APPLICATION NUMBER: 60/099,658
; EARLIER FILING DATE: 1998-09-09
; NUMBER OF SEQ ID NOS: 578
; SOFTWARE: Patent.pm
; SEQ ID NO 198
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1..47
; OTHER INFORMATION: polymorphic fragment 4-4-187
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: polymorphic base A
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 1..23
; OTHER INFORMATION: potential microsequencing oligo 4-4-187.misl
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 25..47
; OTHER INFORMATION: complement potential microsequencing oligo 4-4-187.mis2
US-09-218-207-198

Query Match          1.8%; Score 27; DB 4; Length 47;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1452 AGTAAAAA 1478
Db 21 AGTAAAAA 47

RESULT 32
US-09-328-111-561/c
; Sequence 561, Application US/09328111
```

```
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 561
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(569)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-561

Query Match          1.8%; Score 27; DB 4; Length 569;
Best Local Similarity 100.0%; Pred. No. 0.012;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1452 AGTAAAAA 1478
Db 40 AGTAAAAA 14

RESULT 33
US-08-757-046A-5
; Sequence 5, Application US/08/57046A
; Patent No. 5876995
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; TITLE OF INVENTION: BIOLUMINESCENT ARTICLES OF MANUFACTURE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,046A
; FILING DATE: 11-25-96
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/597,274
; FILING DATE: 02-06-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6680-105B
```

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
;
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 958 base pairs
; TYPE: nucleic acid
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
;
; NAME/KEY: Coding Sequence
; LOCATION: 115...702
; OTHER INFORMATION: apoaeguorin-encoding gene
; PUBLICATION INFORMATION:
; AUTHORS: Inouye et al.
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 82
; PAGES: 3154-3158
; DATE: (1985)
; DOCUMENT NUMBER: PATENT NO.: 5,093,240
;
; US-08-757-046A-5
;
; Query Match 1.8%; Score 27; DB 2; Length 958;
; Best Local Similarity 100.0%; Pred. No. 0.011;
; Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1452 AGTAAAAA...AAAAAAAAAAAAA 1478
; Db 892 AGTAAAAA...AAAAAAAAAAAAA 918
;
; RESULT 34
; US-09-447-208-5
; Sequence 5, Application US/09447208
; Patent No. 6113886
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; TITLE OF INVENTION: BIOLUMINESCENT ARTICLES OF MANUFACTURE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/447,208
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 0909/135,988
; FILING DATE: 08-17-98
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/757,046
; FILING DATE: 11-25-96
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/597,274
; FILING DATE: 02-06-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
;
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24727-105C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-450-8499
;
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 958 base pairs
; TYPE: nucleic acid
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
;
; NAME/KEY: Coding Sequence
; LOCATION: 115...702
; OTHER INFORMATION: apoaeguorin-encoding gene
; PUBLICATION INFORMATION:
; AUTHORS: Inouye et al.
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 82
; PAGES: 3154-3158
; DATE: (1985)
; DOCUMENT NUMBER: PATENT NO.: 5,093,240
;
; US-09-447-208-5
;
; Query Match 1.8%; Score 27; DB 3; Length 958;
; Best Local Similarity 100.0%; Pred. No. 0.011;
; Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1452 AGTAAAAA...AAAAAAAAAAAAA 1478
; Db 892 AGTAAAAA...AAAAAAAAAAAAA 918
;
; RESULT 35
; US-09-135-988-5
; Sequence 5, Application US/09135988
; Patent No. 6152358
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; TITLE OF INVENTION: BIOLUMINESCENT ARTICLES OF MANUFACTURE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/135,988
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/757,046
; FILING DATE: 11-25-96
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/597,274
; FILING DATE: 02-06-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
;

```

```
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24727-105C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-450-8400
; TELEFAX: 619-450-8499
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 115...702
; OTHER INFORMATION: apoaeguorin-encoding gene
; PUBLICATION INFORMATION: PATENT NO.: 5,093,240
; AUTHORS: Inouye et al.
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 82
; PAGES: 3154-3158
; DATE: (1985)
; US-09-135-988-5
;
; Query Match 1.8%; Score 27; DB 3; Length 958;
; Best Local Similarity 100.0%; Pred. No. 0.011;
; Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1452 AGTAAAGAAAAAAGAAAAAAGAAAAA 1478
; Db 892 AGTAAAGAAAAAAGAAAAAAGAAAAA 918
;
; RESULT 36
; US-09-277-716-5
; Sequence 5, Application US/09277716A
; Patent No. 6232107
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROMUNE, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
; CURRENT APPLICATION NUMBER: US/09/277,716A
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/102,939
; EARLIER FILING DATE: 1998-10-01
; EARLIER APPLICATION NUMBER: 60/089,367
; EARLIER FILING DATE: 1998-06-15
; EARLIER APPLICATION NUMBER: 60/079,624
; EARLIER FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 958
; TYPE: DNA
; ORGANISM: Aequorea (luminescent jellyfish)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)..(702)
; FEATURE:
; OTHER INFORMATION: Apoaeguorin-encoding gene
; PUBLICATION INFORMATION:
; PATENT DOCUMENT NUMBER: 5,093,240
; PATENT FILING DATE: 1987-10-08
; PUBLICATION DATE: 1992-03-03
; PUBLICATION INFORMATION:
; AUTHORS: Inouye, S.
;
; ;
; TITLE: Cloning and sequence analysis of cDNA for the luminescent protein aequo
; JOURNAL: Proc. Natl. Acad. Sci. USA
; VOLUME: 82(10)
; PAGES: 3154-3158
; DATE: 1985-05
; TELEFAX: 619-277-716-5
; US-09-277-716-5
;
; Query Match 1.8%; Score 27; DB 4; Length 958;
; Best Local Similarity 100.0%; Pred. No. 0.011;
; Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1452 AGTAAAGAAAAAAGAAAAAAGAAAAA 1478
; Db 892 AGTAAAGAAAAAAGAAAAAAGAAAAA 918
;
; RESULT 37
; US-08-597-274A-5
; Sequence 5, Application US/08597274A
; Patent No. 6247995
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; TITLE OF INVENTION: BIOLUMINESCENT NOVELTY ITEMS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/597,274A
; FILING DATE: 02/06/96
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6680-105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEtical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 115...702
; OTHER INFORMATION: apoaeguorin-encoding gene
; PUBLICATION INFORMATION:
; DOCUMENT NUMBER: 5,093,240
; AUTHORS: Inouye et al.
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 82
; PAGES: 3154-3158
; DATE: (1985)
```

US-08-597-274A-5

Query Match 1.8%; Score 27; DB 4; Length 958;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 AGTAAAGAAAAAAGAAAAAAGAAAAA 1478
|||||
DB 892 AGTAAAGAAAAAAGAAAAAAGAAAAA 918

RESULT 38

US-08-468-819-88
Sequence 88, Application US/08468819
Patent No. 5871723

GENERAL INFORMATION:
APPLICANT: Strieter, Robert M.
APPLICANT: Polverini, Peter J.
APPLICANT: Kunkel, Steven L.
TITLE OF INVENTION: CXG Chemokines as Regulators of
TITLE OF INVENTION: Angiogenesis
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,819
FILING DATE: Concurrently herewith
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: UMIC:003/HYL
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7477
TELEX: N/A

INFORMATION FOR SEQ ID NO: 88:
SEQUENCE CHARACTERISTICS:
LENGTH: 1173 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "DNA"

US-08-468-819-88

Query Match 1.8%; Score 27; DB 2; Length 1173;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 AGTAAAGAAAAAAGAAAAAAGAAAAA 1478
|||||
DB 1143 AGTAAAGAAAAAAGAAAAAAGAAAAA 1169

RESULT 39

US-08-756-387B-1
Sequence 1, Application US/08756387B
Patent No. 5945294

GENERAL INFORMATION:
APPLICANT: Frank, Glenn R.
APPLICANT: Porter, James P.
APPLICANT: Rushlow, Keith E.

APPLICANT: Wassom, Donald L.
TITLE OF INVENTION: Method to Detect IgE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heskia Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/756,387B
FILING DATE: No. 5945294ember 26, 1996

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: DI-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1198 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:

NAME/KEY: CDS
LOCATION: 107..877
US-08-756-387B-1

Query Match 1.8%; Score 27; DB 2; Length 1198;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1452 AGTAAAGAAAAAAGAAAAAAGAAAAA 1478
|||||
DB 1169 AGTAAAGAAAAAAGAAAAAAGAAAAA 1195

RESULT 40

US-08-756-387B-3/c
Sequence 3, Application US/08756387B
Patent No. 5945294

GENERAL INFORMATION:
APPLICANT: Frank, Glenn R.
APPLICANT: Porter, James P.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wassom, Donald L.
TITLE OF INVENTION: Method to Detect IgE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:

ADDRESSEE: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heskia Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:

```
;
; APPLICATION NUMBER: US/08/756.387B
; FILING DATE: NO. 5945294ember 26, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1198 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-756-387B-3

Query Match 1.8%; Score 27; DB 2; Length 1198;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1452 AGTAAAAA 1478
| 30 AGTAAAAA 4

Db 30 AGTAAAAA 4

RESULT 41
US-09-285-873-1
; Sequence 1, Application US/09285873
; Patent No. 6309832
; GENERAL INFORMATION:
; APPLICANT: Frank, Glenn R.
; APPLICANT: Porter, James P.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wassom, Donald L.
; TITLE OF INVENTION: Method to Detect Ige
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/285.873
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/756.387
; FILING DATE: NO. 6309832ember 26, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1198 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
```

```
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..877
; US-09-285-873-1

Query Match 1.8%; Score 27; DB 4; Length 1198;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1452 AGTAAAAA 1478
| 30 AGTAAAAA 1195

Db 1169 AGTAAAAA 1195

RESULT 42
US-09-285-873-3/c
; Sequence 3, Application US/09285873
; Patent No. 6309832
; GENERAL INFORMATION:
; APPLICANT: Frank, Glenn R.
; APPLICANT: Porter, James P.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wassom, Donald L.
; TITLE OF INVENTION: Method to Detect Ige
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/285.873
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/756.387
; FILING DATE: NO. 6309832ember 26, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1198 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-09-285-873-3

Query Match 1.8%; Score 27; DB 4; Length 1198;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1452 AGTAAAAA 1478
| 30 AGTAAAAA 4

Db 30 AGTAAAAA 4

RESULT 43
US-07-913-107-1/c
; Sequence 1, Application US/07913107
```

```

; Patent No. 5382521
; GENERAL INFORMATION:
; APPLICANT: Raz, Avraham
; APPLICANT: Nahi, Ivan R.
; APPLICANT: Otto, Thomas
; APPLICANT: Watanabe, Hideomi
; TITLE OF INVENTION: Method of Determining Metastatic
; TITLE OF INVENTION: Potential of Tumor Cells
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DYKEMA GOSSETT
; STREET: 400 Renaissance Center
; CITY: Detroit
; STATE: MI
; COUNTRY: USA
; ZIP: 48243
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/913,107
; FILING DATE: 19920714
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelly, Robert L.
; REGISTRATION NUMBER: 31,843
; REFERENCE/DOCKET NUMBER: 61,686-016
; TELEPHONE: (313) 540-0849
; TELEFAX: (313) 540-0763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1810 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 177..1145
; US-07-913-107-1
; Query Match 1.8%; Score 27; DB 1; Length 1810;
; Best Local Similarity 100.0%; Pred. No. 0.011;
; Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 1452 AGTAAAGAAAAAAGAAAAAAGAAAAA 1478
| | | | | | | | | | | | | | | | | |
Db 1530 AGTAAAGAAAAAAGAAAAAAGAAAAA 1504

RESULT 45
US-08-281-248-1/c
; Sequence 1, Application US/08281248
; Patent No. 5650500
; GENERAL INFORMATION:
; APPLICANT: Raz, Avraham
; APPLICANT: Nahi, Ivan R.
; APPLICANT: Otto, Thomas
; APPLICANT: Watanabe, Hideomi
; TITLE OF INVENTION: Method of Determining Metastatic
; TITLE OF INVENTION: Potential of Tumor Cells
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DYKEMA GOSSETT
; STREET: 400 Renaissance Center
; CITY: Detroit
; STATE: MI
; COUNTRY: USA
; ZIP: 48243
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/281,248
; FILING DATE: 27-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/913,107
; FILING DATE: 14-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelly, Robert L.
; REGISTRATION NUMBER: 31,843
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,201
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/913,107
; FILING DATE: 14-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelly, Robert L.
; REGISTRATION NUMBER: 31,843

```


REFERENCE/DOCKET NUMBER: 61,686-016
TELEPHONE: (313) 540-0849
TELEFAX: (313) 540-0763
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1810 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 177..1145
US-08-281-248-1

Query Match 1.8%; Score 27; DB 1; Length 1810;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1452 AGTAAAAA 1478
Db 1530 AGTAAAAA 1504

Search completed: November 5, 2002, 13:53:45
Job time : 248.249 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 13:04:56 ; Search time 1264.61 Seconds
(without alignments)
15774.429 Million cell updates/sec

Title: US-09-805-311-7
Perfect score: 1478
Sequence: 1 cgaccacgcgtccgggaaa.....aaaaaaaaaaaaaaaaaaaaa 1478

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST.*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	570	38.6	586	9 AI881599	AI881599 606068G09
2	452	30.6	553	9 AW562789	AW562789 660065H06
3	411	27.8	475	9 AW562517	AW562517 660065H06
4	388	26.3	474	9 AW559173	AW559173 660065H06
5	347	23.5	467	9 BE186786	BE186786 946012C08
6	301	20.4	363	9 AW562518	AW562518 660065H06
7	265	17.9	901	10 BG837708	BG837708 zml0_01f0
8	259	17.5	550	10 BE639422	BE639422 946033A02
9	240	16.2	554	10 BE639421	BE639421 946033A02
10	225	15.2	225	9 AW562788	AW562788 660065H06
11	223	15.1	456	9 AW065689	AW065689 ag91f12.x
12	204	13.8	532	9 AW000375	AW000375 614014D03
13	191	12.9	414	9 AW288784	AW288784 707010C02
14	184	12.4	376	9 AI065546	AI065546 ag88e02.x
15	184	12.4	470	9 AI861468	AI861468 614014D03
16	173	11.7	470	9 AW288831	AW288831 707010F11
17	148	10.0	470	9 AW288831	AW288831 707010F11

C	18	134	9.1	553	9	AI834484	AI834484 606068G09
19	103	7.0	232	10	BF727781	BF727781 1000052F1	
C	20	68	4.6	105	9	AI932215	AI932215 618029H11
21	68	4.6	126	9	AW147048	AW147048 707010C02	
C	22	65	4.4	126	9	AW147048	AW147048 707010C02
C	23	57	3.9	225	9	AI947478	AI947478 614047B01
24	44	3.0	253	10	D43467	D43467 D43467 Rice	
25	38	2.6	454	9	AV915644	AV915644 AV915644	
26	38	2.6	648	9	AV913663	AV913663 AV913663	
27	35	2.4	539	10	BF098320	BF098320 EST428761	
28	35	2.4	592	9	AI780966	AI780966 EST261845	
29	31	2.1	143	9	AI271296	AI271296 qw68q01.x	
30	31	2.1	152	9	AI308412	AI308412 tb43e12.x	
31	31	2.1	157	9	AI305700	AI305700 qw47e05.x	
32	31	2.1	163	9	AI250739	AI250739 qx23d02.x	
33	31	2.1	179	9	AI266300	AI266300 qp69c03.x	
C	34	31	2.1	253	9	AI825335	AI825335 wdi17b09.x
35	31	2.1	260	9	AW706512	AW706512 sj57d02.y	
C	36	31	2.1	326	9	AI078141	AI078141 oz30b03.x
C	37	31	2.1	366	9	AI394973	AI394973 MA001960
38	31	2.1	408	9	AU165690	AU165690 AU165690	
C	39	31	2.1	444	9	AI017545	AI017545 ou35c07.x
C	40	31	2.1	479	9	AW019727	AW019727 fd56d01.x
41	31	2.1	573	9	AA454191	AA454191 zx48b10.f	
C	42	31	2.1	640	9	AW071753	AW071753 ws55c02.x
C	43	31	2.1	681	10	BM078557	BM078557 MEST121-C
C	44	30	2.0	120	9	AA481494	AA481494 aa34a11.s
C	45	30	2.0	137	9	AI968696	AI968696 wt91f09.x

ALIGNMENTS

AI881599 586 bp mRNA linear EST 02-FEB-2000
606068G09.yl 606 - Ear tissue cDNA library from Schmidt lab Zea
mays cDNA, mRNA sequence.
AI881599.1 GI:5566733
EST.
Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 586)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 606068 row: G column: 09.

FEATURES

source
Location/Qualifiers
1..586
/organism="Zea mays"
/cultivar="Chilo43"
/db_xref="taxon:4577"
/clone_lib="606 - Ear tissue cDNA library from Schmidt
lab"
/tissue_type="mixed"
/dev_stage="ear length from 0.5 cm - 2.0 cm"
/lab_host="XLOLR (Stratagene)"
/note="Organ: immature ear; Vector: pBK-CWV; Site_1: EcoRI
; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt
lab"

BASE COUNT 186 a 132 c 159 g 109 t

JOURNAL
 COMMENT
 University
 Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 660065 row: H column: 06.
 Location/Qualifiers
 1..475
 /organism="Zea mays"
 /cultivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="660 - Mixed stages of anther and pollen"
 /tissue_type="whole premeiotic anthers to pollen shed"
 /dev_stage="premeiotic anthers to pollen shed"
 /lab_host="XLOLR"
 /note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
 Site_2: XhoI; Anther and pollen cDNA library.
 Directionally sequenced with 5' end at the EcoRI site.
 Created by Amie Franklin."

BASE COUNT 124 a 110 c 94 g 147 t
 ORIGIN
 Query Match 27.8%; Score 411; DB 9; Length 475;
 Best Local Similarity 100.0%; Pred. No. 1.5e-79; Indels 0; Gaps 0;
 Matches 411; Conservative 0; Mismatches 0;

QY 991 ACTGCACCTGATGAGGAGGCTCATAGTTTCTCGTAAAGATAATGTTTCAATGAA 1050
 DB 475 ACTGCACCTGATGAGGAGGCTCATAGTTTCTCGTAAAGATAATGTTTCAATGAA 416
 QY 1051 GATCGGGTGACAAAGGCCATAGAGAAGATCAATCTGCCAAGAAATAATCGTCGAAGGA 1110
 DB 415 GATCGGGTGACAAAGGCCATAGAGAAGATCAATCTGCCAAGAAATAATCGTCGAAGGA 356
 QY 1111 AGACTCGAGTCTCTTTTCAAGCCAACTGCCACCATCATCAGCAGCGCTAAAGCGAAGGAG 1170
 DB 355 AGACTCGAGTCTCTTTTCAAGCCAACTGCCACCATCATCAGCAGCGCTAAAGCGAAGGAG 296
 QY 1171 ACTTCGGATAAACAAGCAAGCGAGCTGCCAACAAGAAACAAGGCTGCTGGAAGGAAG 1230
 DB 295 ACTTCGGATAAACAAGCAAGCGAGCTGCCAACAAGAAACAAGGCTGCTGGAAGGAAG 236
 QY 1231 AATAATCTTGATGCTTGATGTACAACACTACGACTACGAAACAGCGGTGGCATGATCAC 1290
 DB 235 AATAATCTTGATGCTTGATGTACAACACTACGACTACGAAACAGCGGTGGCATGATCAC 176
 QY 1291 TTCCGCTAGATTATTAACTCCCTGTTTAACTCAGACCTTTTGGTCAAAAGTTTGCCCATG 1350
 DB 175 TTCCGCTAGATTATTAACTCCCTGTTTAACTCAGACCTTTTGGTCAAAAGTTTGCCCATG 116
 QY 1351 TTTCAGCTGGGGTAAGTTAGTTAGTTGTTTGAAGAGATTGGTGTAACCAAGTA 1401
 DB 115 TTTCAGCTGGGGTAAGTTAGTTAGTTGTTTGAAGAGATTGGTGTAACCAAGTA 65

RESULT 4
 AW559173/c 474 bp mRNA linear EST 07-MAR-2000
 LOCUS BE186786/c
 DEFINITION 660065H06.x1 660 - Mixed stages of anther and pollen Zea mays cDNA,
 mRNA sequence.
 ACCESSION AW559173
 VERSION AW559173.1 GI:7204640
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 REFERENCE 1 (bases 1 to 474)

AUTHORS Walbot V.
 TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
 JOURNAL University
 COMMENT Unpublished (1999)
 Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 660065 row: H column: 06.
 Location/Qualifiers
 1..474
 /organism="Zea mays"
 /cultivar="Ohio43"
 /db_xref="taxon:4577"
 /clone_lib="660 - Mixed stages of anther and pollen"
 /tissue_type="whole premeiotic anthers to pollen shed"
 /dev_stage="premeiotic anthers to pollen shed"
 /lab_host="XLOLR"
 /note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
 Site_2: XhoI; Anther and pollen cDNA library.
 Directionally sequenced with 5' end at the EcoRI site.
 Created by Amie Franklin."

BASE COUNT 113 a 114 c 92 g 154 t
 ORIGIN
 Query Match 26.3%; Score 388; DB 9; Length 474;
 Best Local Similarity 99.8%; Pred. No. 1.3e-74; Indels 0; Gaps 0;
 Matches 438; Conservative 0; Mismatches 1;

QY 967 TTGGATATTCCTGAGCTAAAATGGACTGCACCTGATGAGGAGGCTCATAGATTTCCCTG 1026
 DB 439 TTGGATATTCCTGAGCTAAAATGGACTGCACCTGATGAGGAGGCTCATAGATTTCCCTG 380
 QY 1027 GTAAGAATAATGTTTCATGAGATCGGGTGACAAAGGCCATAGAGAGATCAATCT 1086
 DB 379 GTAAGAATAATGTTTCATGAGATCGGGTGACAAAGGCCATAGAGAGATCAATCT 320
 QY 1087 GCCAAGAATAATCGTCGCAAGGAAGACTCGAGTCTCTTTTCAAGCCAACTGCCACCACA 1146
 DB 319 GCCAAGAATAATCGTCGCAAGGAAGACTCGAGTCTCTTTTCAAGCCAACTGCCACCACA 260
 QY 1147 TCAGCACCGCTTAACCGAGGAGACTTCGGATTAACCAAGCAAGCAGCAGCTCGGAACAAG 1206
 DB 259 TCAGCACCGCTTAACCGAGGAGACTTCGGATTAACCAAGCAAGCAGCAGCTCGGAACAAG 200
 QY 1207 AAAACAAGGCTGGTGGAAAGAGAAAATAATCTTGGATGCTTGTATGATACAACTACGACTA 1265
 DB 199 AAAACAAGGCTGGTGGAAAGAGAAAATAATCTTGGATGCTTGTATGATACAACTACGACTA 140
 QY 1267 CGAAAGCAGCGGTGGCATGATCACTTCGCCCTAGATTAATTAACCTCCCTGTTTAACCTAG 1326
 DB 139 CGAAAGCAGCGGTGGCATGATCACTTCGCCCTAGATTAATTAACCTCCCTGTTTAACCTAG 80
 QY 1327 ACCTTTGGTGAAGTTTGCCCATGTTTCAAGCTGGGGTAAGTTAGTTGTTGAAGAGA 1386
 DB 79 ACCTTTGGTGAAGTTTGCCCATGTTTCAAGCTGGGGTAAGTTAGTTGTTGAAGAGA 20
 QY 1387 TTGGTGTACCAAGTAACAA 1405
 DB 19 TTGGTGTACCAAGTAACAA 1

RESULT 5
 BE186786/c 467 bp mRNA linear EST 22-JUN-2000
 LOCUS BE186786
 DEFINITION 946012C08.X1 946 - tassal primordium prepared by Schmidt lab Zea
 mays cDNA, mRNA sequence.
 ACCESSION BE186786
 VERSION BE186786.1 GI:8665970
 KEYWORDS EST.

SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
TITLE 1 (bases 1 to 467)
JOURNAL Zea mays.
COMMENT Zea mays
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946012 row: C column: 08.

FEATURES
source
1..467
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassels primordium prepared by Schmidt
lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XLOLR"
/note="Organ: tassels; Vector: HybridZAP; Site_1: EcoRI;
Site_2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybridZAP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."
BASE COUNT 108 a 110 c 96 g 153 t
ORIGIN

Query Match 23.58; Score 347; DB 9; Length 467;
Best Local Similarity 99.7%; Pred. No. 8.le-66;
Matches 397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 898 AGATATCAAAATTCCTGAGGACTGGCCCTTACCAAGAAGCTCGAGCTGTGTTCAAGGAGCCT 957
DB 467 AGATATCAAAATTCCTGAGGACTGGCCCTTACCAAGAAGCTCGAGCTGTGTTCAAGGAGCCT 408
QY 958 ATGTCACTTGGATATTCCTGAGCTAAATGGACTGCACCTGATGAGAGGCTCTCAT 1017
DB 407 AATGTCACTTGGATATTCCTGAGCTAAATGGACTGCACCTGATGAGAGGCTCTCAT 348
QY 1018 AGTTTCTGGTAAAGATAATGTTTCAATGAAGATCGGCTGACAAAGGCCATAGAGAAG 1077
DB 347 AGTTTCTGGTAAAGATAATGTTTCAATGAAGATCGGCTGACAAAGGCCATAGAGAAG 288
QY 1078 ATCAATCTGCCACATAAATCTGCCAGGAGACTCGAGTCTTTTCAAGCCACT 1137
DB 287 ATCAATCTGCCACATAAATCTGCCAGGAGACTCGAGTCTTTTCAAGCCACT 228
QY 1138 GCCACCATCAGCACCGCTAAACCGAAGGAGACTTCGATATAACAAAGCAGCT 1197
DB 227 GCCACCATCAGCACCGCTAAACCGAAGGAGACTTCGATATAACAAAGCAGCT 168
QY 1198 GCGACACAGAAACAAAGGCTGGTGGAAAGAAATAATCTTGGATGCTTGATGTACAA 1257
DB 167 GCGACACAGAAACAAAGGCTGGTGGAAAGAAATAATCTTGGATGCTTGATGTACAA 108
QY 1258 CTACGACTACGAAGCAGCGGTGGCATGATCACTTCGC 1295
DB 107 CTACGACTACGAAGCAGCGGTGGCATGATCACTTCGC 70

RESULT 6
AW562518/c
LOCUS AW562518 363 bp mRNA linear EST 10-MAR-2000

DEFINITION 660065H06.x3 660 - Mixed stages of anther and pollen Zea mays cDNA,
mRNA sequence.
ACCESSION AW562518
VERSION AW562518.1 GI:7216396
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
TITLE 1 (bases 1 to 363)
JOURNAL Zea mays.
COMMENT Zea mays
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660065 row: H column: 06.

FEATURES
source
1..363
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premeiotic anthers to pollen shed"
/dev_stage="premeiotic anthers to pollen shed"
/lab_host="XLOLR"
/note="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 96 a 82 c 74 g 111 t
ORIGIN

Query Match 20.4%; Score 301; DB 9; Length 363;
Best Local Similarity 100.0%; Pred. No. 7.5e-56;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 CCGCTAAACCGAAGGAGACTTCGGATAAAACAAAGCAGCTCGCAACAGAAACA 1212
DB 301 CCGCTAAACCGAAGGAGACTTCGGATAAAACAAAGCAGCTCGCAACAGAAACA 242
QY 1213 AAGGCTGGTGGAAAGAAATAATCTTCGATGCTTGATGTACAACTACGACTACGAAAG 1272
DB 241 AAGGCTGGTGGAAAGAAATAATCTTCGATGCTTGATGTACAACTACGACTACGAAAG 182
QY 1273 CAGCGTGGCATGATCACTTCGCTAGATTATTAACTCCCTGTTTAACTCAGACCTTT 1332
DB 181 CAGCGTGGCATGATCACTTCGCTAGATTATTAACTCCCTGTTTAACTCAGACCTTT 122
QY 1333 GGTAAAGTTTGGCCATGTTTCAAGCTGGGGTAAGTTAGTTGTTGTTGAAGAGATTGGTG 1392
DB 121 GGTAAAGTTTGGCCATGTTTCAAGCTGGGGTAAGTTAGTTGTTGTTGAAGAGATTGGTG 62
QY 1393 TACCAAGTAACAAACCTATATCGTGTGTTTTTACTTCTGCTCTTTTGAAGTATGTATGCCA 1452
DB 61 TACCAAGTAACAAACCTATATCGTGTGTTTTTACTTCTGCTCTTTTGAAGTATGTATGCCA 2
QY 1453 G 1453
DB 1 G 1

RESULT 7
BG837708
LOCUS BG837708 901 bp mRNA linear EST 25-MAY-2001
DEFINITION Zm10_01f08_A Zm10_AAPC_ECORC_Fusarium_graminearum_corn_silk Zea
mays cDNA clone Zm10_01f08, mRNA sequence.

```

ACCESSION      BG837708
VERSION        BG837708.1  GI:14204031
KEYWORDS       Zea mays.
SOURCE         Zea mays.
ORGANISM       Zea mays.

REFERENCE      1 (bases 1 to 901)
AUTHORS        Harris,L.J., Balcerzak,M., Allard,S., Saparno,A., Couroux,P., De
               Moors,A., Hattori,J.I., Ouellet,F., Robert,L.S., Singh,J.A, Sprrott
               ,D. and Tinker,N.A.
TITLE          Expressed Sequence Tags from Maize Silk Six Hours After Silk
               Channel Inoculation with Fusarium graminearum
JOURNAL        Unpublished (2001)
COMMENT        Contact: Harris, Linda J.
               Eastern Cereal and Oilseed Research Centre
               Agriculture and Agri-food Canada
               Bldg. 21, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
               CANADA
               Tel: (613) 759-1314
               Fax: (613) 759-6566
               Email: harrisl@eem.agr.ca.

FEATURES       Location/Qualifiers
source         1..901
               /organism="Zea mays"
               /cultivar="CO388"
               /db_xref="taxon:4577"
               /clone_lib="Zm10_01f08"
               /tissue_type="Silk"
               /dev_stages="4-5 days post-silk emergence"
               /note="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI;
               Site_2: XhoI; Field-grown corn was silk channel-inoculated
               in the morning (~10 am) with 1 ml of a macroconidial
               suspension (500,000 spores/ml) of Fusarium graminearum and
               silk channels were collected and immediately frozen in
               liquid nitrogen 6 hours later. RNA was extracted from
               silk tissue between 1 cm below and above the inoculation
               point in the silk channel. RNA from five silk channels was
               pooled."
BASE COUNT     267 a 174 c 208 g 249 t 3 others
ORIGIN
Query Match    17.9%; Score 265; DB 10; Length 901;
Best Local Similarity 99.7%; Pred. No. 1.7e-48;
Matches 315; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 967 TTGGATATTCCTGAGCTAAATGGAGTGCACCTGATGAGGAGGCTCTCATAGCTTTCCCTG 1026
Db 353 TTGGATATTCCTGAGCTAAATGGAGTGCACCTGATGAGGAGGCTCTCATAGCTTTCCCTG 412
Qy 1027 GTAAGATTAATGGTTTCAATGAAGATCGGGTGACAAAGGCCATAGAGAAGATCAAAATCT 1086
Db 413 GTAAGATTAATGGTTTCAATGAAGATCGGGTGACAAAGGCCATAGAGAAGATCAAAATCT 472
Qy 1087 GCCAAGAAATAATGTCGCAAGGAAGACTCGAGTCTCTTTTCAAGCCAACTGCCACACA 1146
Db 473 GCCAAGAAATAATGTCGCAAGGAAGACTCGAGTCTCTTTTCAAGCCAACTGCCACACA 532
Qy 1147 TCAGCACCGCTAAACGGAGGAGACTTCGGATTAACCAAGCAAGGCGAGCTGGCAACAAG 1206
Db 533 TCAGCACCGCTAAACGGAGGAGACTTCGGATTAACCAAGCAAGGCGAGCTGGCAACAAG 592
Qy 1207 AAAACAAAGGCTGGTGAAGGAAGAAATAATCTTGGATGCTTGGATGACAACTACGACTA 1266
Db 593 AAAACAAAGGCTGGTGAAGGAAGAAATAATCTTGGATGCTTGGATGACAACTACGACTA 652
Qy 1267 CGAAAGCAGCGGTGC 1282
Db 653 CGAAAGCAGCGGTGC 668

```

```

RESULT 8
LOCUS      BE639422
DEFINITION 946033A02.y2 946 - tassell primordium prepared by Schmidt lab Zea
mays cDNA, mRNA sequence.
ACCESSION  BE639422
VERSION     BE639422
KEYWORDS    EST.
SOURCE      BE639422.1  GI:9952839
ORGANISM    Zea mays.
            Zea mays.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
            1 (bases 1 to 550)
            Walbot,V.
            Maize ESTs from various cDNA libraries sequenced at Stanford
            University
            Unpublished (1999)
            Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 946033 row: A column: 02.
            Location/Qualifiers
            1..550
            /organism="Zea mays"
            /cultivar="OH43"
            /db_xref="taxon:4577"
            /clone_lib="946 - tassell primordium prepared by Schmidt
            lab"
            /tissue_type="tassels"
            /dev_stages="just after the transition from vegetative to
            inflorescence development"
            /lab_host="XL0LR"
            /note="Organ: tassels; Vector: HybriZAP; Site_1: EcoRI;
            Site_2: XhoI; George Chuck dissected immature tassels
            between 1mm and 3mm. Sharon Stanfield prepared the cDNA
            library in HybriZAP. Sample insert size range was 350 bp
            to 3 kb with a 1 kb average."
BASE COUNT     177 a 116 c 149 g 108 t
ORIGIN
Query Match    17.5%; Score 259; DB 10; Length 550;
Best Local Similarity 99.4%; Pred. No. 5.2e-47;
Matches 359; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 246 CATGGAAACTCTCACAATGAAGCTGGTGAAGTCACTAGTCAATTCGAAGAATGTTCAA 305
Db 190 CATGGAAACTCTCACAATGAAGCTGGTGAAGTCACTAGTCAATTCGAAGAATGTTCAA 249
Qy 306 CCGSACAATAAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGTGSCAAGCC 365
Db 250 CCGSACAATAAGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGTGSCAAGCC 309
Qy 366 TCCTGATATCAAGAAACAAGAACTTGCTAAAGATACCTCAAAAAGAGATGATGCAACCA 425
Db 310 TCCTGATATCAAGAAACAAGAACTTGCTAAAGATACCTCAAAAAGAGATGATGCAACCA 369
Qy 426 AGATCTGACTGAGGCGAGTAGAGGATAAAGATCGGATTTGAAAATTTGAGCAAGAG 485
Db 370 AGATCTGACTGAGGCGAGTAGAGGATAAAGATCGGATTTGAAAATTTGAGCAAGAG 429
Qy 486 GACTGTAAAGGTCACAAAGCAACACACAGCATGTAACGACTATTAAGACTTATGGG 545
Db 430 GACTGTAAAGGTCACAAAGCAACACACAGCATGTAACGCGCTATTAAAGACTTATGGG 489
Qy 546 GGTTCCTGTTGTAGAGGCACCTTCTGAAGCAGAGAGAATGTGACGCCCTTTGCATAAA 605
Db 490 GGTTCCTGTTGTAGAGGCACCTTCTGAAGCAGAGAGAATGTGACGCCCTTTGCATAAA 549

```

```

QY 606 C 606
Db 550 C 550

RESULT 9
BE639421/C
LOCUS
DEFINITION 554 bp mRNA linear EST 30-AUG-2000
946033A02.x2 946 - tassels primordium prepared by Schmidt lab Zea
mays cDNA, mRNA sequence.
ACCESSION BE639421
VERSION BE639421
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946033 row: A column: 02.

FEATURES
source
1..554
/organism="Zea mays"
/cultivar="OH43"
/db_xref="taxon:4577"
/clone_lib="946 - tassels primordium prepared by Schmidt
lab"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XLOLR"
/notes="Organ: tassels; Vector: HybridZAP; Site_1: EcoRI;
Site_2: XhoI; George Chuck dissected immature tassels
between Imm and 3mm. Sharon Stanfield prepared the cDNA
library in HybridZAP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."
BASE COUNT 145 a 131 c 110 g 168 t

Query Match 16.2%; Score 240; DB 10; Length 554;
Best Local Similarity 99.7%; Pred. No. 6.1e-43;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 992 CTCACCTGATGAGGAGGCTCTCATAAGTTTCTCGTAAAGATAATGGTTTCAATCAAG 1051
|||||
Db 470 CTCACCTGATGAGGAGGCTCTCATAAGTTTCTCGTAAAGATAATGGTTTCAACGAAG 411
|||||

QY 1052 ATCGGTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGATAAATCGTCGAAGGAA 1111
|||||
Db 410 ATCGGTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGATAAATCGTCGAAGGAA 351
|||||

QY 1112 GACTCGAGTCCTTTTTCAGCCCAACTGCCACCACATCAGCCGCTAAACGGAAGGAGA 1171
|||||
Db 350 GACTCGAGTCCTTTTTCAGCCCAACTGCCACCACATCAGCCGCTAAACGGAAGGAGA 291
|||||

QY 1172 CTTCCGATAAAACACGAGGAGCTGCCAACAACAAACAGCTCGTGGAAAGAGA 1231
|||||
Db 290 CTTCCGATAAAACACGAGGAGCTGCCAACAACAAACAGCTCGTGGAAAGAGA 231
|||||

QY 1232 AATAATCTTGATGCTTGATGTACAACTACGACTACGAAAGCAGCGGTGGC 1282
|||||

```

```

Db 230 AATAATCTTGATGCTTGATGTACAACTACGACTACGAAAGCAGCGGTGGC 180

RESULT 10
AW562788/C
LOCUS
DEFINITION 225 bp mRNA linear EST 10-MAR-2000
660065H06.x5 660 - Mixed stages of anther and pollen Zea mays cDNA,
mRNA sequence.
ACCESSION AW562788
VERSION AW562788
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.

REFERENCE
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 660065 row: H column: 06.

FEATURES
source
1..225
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="660 - Mixed stages of anther and pollen"
/tissue_type="whole premiotic anthers to pollen shed"
/dev_stage="premiotic anthers to pollen shed"
/lab_host="XLOLR"
/notes="Organ: anthers; Vector: Lambda Zap; Site_1: EcoRI;
Site_2: XhoI; Anther and pollen cDNA library.
Directionally sequenced with 5' end at the EcoRI site.
Created by Amie Franklin."
BASE COUNT 40 a 53 c 45 g 87 t

Query Match 15.2%; Score 225; DB 9; Length 225;
Best Local Similarity 100.0%; Pred. No. 2.3e-39;
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1010 GTCTCATAAGTTTCTCGTAAAGATAATGGTTTCAATGAAGATCGGTGACAAAGGCCA 1069
|||||
Db 225 GTCTCATAAGTTTCTCGTAAAGATAATGGTTTCAATGAAGATCGGTGACAAAGGCCA 166
|||||

QY 1070 TAGAGAAGATCAAAATCTGCCAAGATAAATCGTGCAGGAAGACTCGAGTCCTTTTCA 1129
|||||
Db 165 TAGAGAAGATCAAAATCTGCCAAGATAAATCGTGCAGGAAGACTCGAGTCCTTTTCA 106
|||||

QY 1130 AGCCAACTGCCACCACATCAGCACCGCTAAACCGGAAGAGACTTCGGATAAAACAAAGCA 1189
|||||
Db 105 AGCCAACTGCCACCACATCAGCACCGCTAAACCGGAAGAGACTTCGGATAAAACAAAGCA 46
|||||

QY 1190 AGCGAGCTCGCAACGAAGAAACAAAGCGCTGGTGGNAAGAGAAAT 1234
|||||
Db 45 AGCGAGCTCGCAACGAAGAAACAAAGCGCTGGTGGNAAGAGAAAT 1

RESULT 11
AI065689/C
LOCUS
DEFINITION 456 bp mRNA linear EST 24-JUL-1998
ag91f12.x1 maize inflorescence immature ear library Zea mays cDNA
clone ag91f12 3', mRNA sequence.
ACCESSION AI065689
VERSION AI065689

```


Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707010 row: C column: 02.

FEATURES
Source
Location/Qualifiers
1..414
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf; Vector: pGAD10; Site_1: EcoRI; cDNA library from fully differentiated maize tissues from an active Mutator plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk, husk, root, leaf). Unidirectionally cloned."
BASE COUNT 120 a 82 c 101 g 109 t 2 others
ORIGIN

Query Match 12.9%; Score 191; DB 9; Length 414;
Best Local Similarity 99.6%; Pred. No. 2.6e-32;
Matches 241; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 661 CCACGGTCTCTTCGTCATTTAATGATCCCAAGTTCACAGAAATACCTGTGATGGAATTT 720
|||||
Db 6 CCACGGTCTCTTCGTCATTTAATGATCCCAAGTTCACAGAAATACCTGTGATGGAATTT 65
|||||

Qy 721 GATGTTGCCAAGGTTTGGAGGAGCTTGAATCACCATGGACCATGTCATTTGTTGTC 780
|||||
Db 66 GATGTTGCCAAGGTTTGGAGGAGCTTGAATCACCATGGACCATGTCATTTGTTGTC 125
|||||

Qy 781 ATCTCTGTGGATGTGACTATTGTGATAGCATCAAGGTATCGGGGCAACAGCTGTG 840
|||||
Db 126 ATCTCTGTGGATGTGACTATTGTGATAGCATCAAGGTATCGGGGCAACAGCTGTG 185
|||||

Qy 841 AAATTTATTCGTAACATGGTCCATAGAAAGCATCTTTGGAGAATCTTTAATAAGACAGA 900
|||||
Db 186 AAATTTATTCGTAACATGGTCCATAGAAAGCATCTTTGGAGAATCTTTAATAAGACAGA 245
|||||

Qy 901 TA 902
||
Db 246 TA 247

RESULT 14
AI065546/c
LOCUS
DEFINITION 376 bp mRNA linear EST 24-JUL-1998
ag88e02.x1 maize inflorescence immature ear library Zea mays cDNA
clone ag88e02 3', mRNA sequence.

ACCESSION
AI065546
VERSION
AI065546.1 GI:3340953
KEYWORDS
EST.
SOURCE
Zea mays.
ORGANISM
Zea mays.

REFERENCE
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
1 (bases 1 to 376)
Clade; Panicoideae; Andropogoneae; Zea.

AUTHORS
Schutz, K., de la Bastide, M., Gnoj, L., Habermann, K., Huang, E. N., Parnell, L. D., Dedhia, N., Martienssen, R. and McCombie, W. R.
Expressed sequence tags from Zea mays
Unpublished (1998)
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: ag88 row: e column: 02

TITLE
JOURNAL
COMMENT

Seq primer: M13 forward universal -21
High quality sequence stop: 376.

FEATURES
Source
Location/Qualifiers
1..376
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="ag88e02"
/sex="female"
/tissue_type="immature ear"
/note="Vector: pLUPEScript SK+ (X52325); Site_1: XhoI; Site_2: EcoRI; This library is described in Schmidt, Hake, et al., (1993) Plant Cell 5:729-737. cDNAs are directionally cloned into the XhoI and EcoRI sites; XhoI is near the polyA tail. Most reads from this library are 3' in direction. Additional information on this library as well as ftp access to all sequences can be found at http://www.cshl.org/maizegenome"

BASE COUNT 79 a 65 g 131 t 2 others
ORIGIN

Query Match 12.4%; Score 184; DB 9; Length 376;
Best Local Similarity 99.6%; Pred. No. 8.9e-31;
Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1048 GAAGATCGGGTGCACAAAGCCATAGAGAGATCAAAATCTGCCAAGATAAATCGTCGCA 1107
|||||
Db 356 GAAGATCGGGTGCACAAAGCCATAGAGAGATCAAAATCTGCCAAGATAAATCGTCGCA 297
|||||

Qy 1108 GGAAGACTCGAGTCTCTTTTCAAGCCAACTGCCACCACATCAGCACCGCTAAACCGGAAG 1167
|||||
Db 296 GGAAGACTCGAGTCTCTTTTCAAGCCAACTGCCACCACATCAGCACCGCTAAACCGGAAG 237
|||||

Qy 1168 GAGACTTCGGATAAAACGAAGCAGCTCGCAACGAAGAAACAAAGCGTGGTGAAG 1227
|||||
Db 236 GAGACTTCGGATAAAACGAAGCAGCTCGCAACGAAGAAACAAAGCGTGGTGAAG 177
|||||

Qy 1228 AAGAATAATCTTGGATCTTGATGTACACTACGACTACGAAAGCGCGTGGC 1282
|||||
Db 176 AAGAATAATCTTGGATCTTGATGTACACTACGACTACGAAAGCGCGTGGC 122
|||||

RESULT 15
AI861468/c
LOCUS
DEFINITION 470 bp mRNA linear EST 19-JUL-1999
614014D03.x1 614 - root cDNA library from Walbot Lab Zea mays cDNA, mRNA sequence.

ACCESSION
AI861468
VERSION
AI861468.1 GI:5525575
KEYWORDS
EST.
SOURCE
Zea mays.
ORGANISM
Zea mays.

REFERENCE
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
1 (bases 1 to 470)
Clade; Panicoideae; Andropogoneae; Zea.

AUTHORS
Walbot, V.
Maize ESTs from various cDNA libraries sequenced at Stanford University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614014 row: D column: 03.
Location/Qualifiers
1..470
/organism="Zea mays"
/cultivar="W23"

TITLE
JOURNAL
COMMENT

```

/db_xref="taxon:4577"
/clone_lib="614 - root cDNA library from Walbot Lab"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XL0LR"
/Note="Organ: root; Vector: pBlueScriptII SK+; Site_1: EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot Lab (LM)"
BASE COUNT      112 a      93 g      152 t
ORIGIN

Query Match      12.4%; Score 184; DB 9; Length 470;
Best Local Similarity 99.6%; Pred. No. 7.2e-31;
Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1048 GAAGATCGGGTGACAAAGCCCATAGAGAAGATCAAAATCTGCCAAGATAAATCGTCGCAA 1107
Db 371 GAAGATCGGGTGACAAAGCCCATAGAGAAGATCAAAATCTGCCAAGATAAATCGTCGCAA 312
QY 1108 GGAAGACTCGAGTCCTTTTCAAGCCCACTGCCACCATCAGCACCGCTAAACCGGAAG 1167
Db 311 GGAAGACTCGAGTCCTTTTCAAGCCCACTGCCACCATCAGCACCGCTAAACCGGAAG 252
QY 1168 GAGACTTCGGATAAACAAGCAAGCGAGCTCGGAACAAGAAACAACAAGGCTGTGGAAG 1227
Db 251 GAGACTTCGGATAAACAAGCAAGCGAGCTCGGAACAAGAAACAACAAGGCTGTGGAAG 192
QY 1228 AGAATATATCTTGATGCTTGTATCACTAGCACTAGCAAGACGCGGTGGC 1282
Db 191 AGAATATATCTTGATGCTTGTATCACTAGCACTAGCAAGACGCGGTGGC 137

RESULT 16
AW288831/c
LOCUS      AW288831      470 bp      mRNA      linear      EST 16-JAN-2000
DEFINITION      mays cDNA, mRNA sequence.
ACCESSION      AW288831
VERSION      AW288831.1      GI:6695753
KEYWORDS      EST.
SOURCE      Zea mays.
ORGANISM      Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 470)
Walbot.V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707010 row: F column: 11.
FEATURES
source
1..470
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK
)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/Note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

BASE COUNT      102 a      110 c      108 g      150 t
ORIGIN

Query Match      11.7%; Score 173; DB 9; Length 470;
Best Local Similarity 99.3%; Pred. No. 1.7e-28;
Matches 273; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 974 TTCTGAGCTAAAATGACTGCACCTGATGAGGAGGCTCTCATAGTTTCTCTGGTAAAG 1033
Db 429 TTCTGAGCTAAAATGACTGCACCTGATGAGGAGGCTCTCATAGTTTCTCTGGTAAAG 370
QY 1034 ATAATGTTTCAATGAAGTCGGGTGACAAAGGCGCATAGAGAAGATCAAAATCTGCCAAGA 1093
Db 369 ATAATGTTTCAACGAAGATCGGGTGACAAAGGCGCATAGAGAAGATCAAAATCTGCCAAGA 310
QY 1094 ATAAATCGTCGCAAGAACACTCGAGTCCTTTTCAAGCCCACTGCCACCATCAGCAC 1153
Db 309 ATAAATCGTCGCAAGAACACTCGAGTCCTTTTCAAGCCCACTGCCACCATCAGCGC 250
QY 1154 CGCTAAACGGGAAGAGACTTCGGATATAAACCAAGCAAGCGAGCTGCCAACAAGAAACAA 1213
Db 249 CGCTAAACGGGAAGAGACTTCGGATATAAACCAAGCAAGCGAGCTGCCAACAAGAAACAA 190
QY 1214 AGCTCGTGGAAAGAAATAAATCTTGGATGCTT 1248
Db 189 AGCTCGTGGAAAGAAATAAATCTTGGATGCTT 155

RESULT 17
AW288831
LOCUS      AW288831      470 bp      mRNA      linear      EST 16-JAN-2000
DEFINITION      mays cDNA, mRNA sequence.
ACCESSION      AW288831
VERSION      AW288831.1      GI:6695753
KEYWORDS      EST.
SOURCE      Zea mays.
ORGANISM      Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 470)
Walbot.V.
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 707010 row: F column: 11.
FEATURES
source
1..470
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="707 - Mixed adult tissues from Walbot lab (SK
)"
/tissue_type="tassel, kernel, silk, husk, root, leaf"
/dev_stage="adult"
/lab_host="DH10B"
/Note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."

BASE COUNT      102 a      110 c      108 g      150 t
ORIGIN

Query Match      10.0%; Score 148; DB 9; Length 470;

```

```

Best Local Similarity 100.0%; Pred. No. 3.8e-23;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 729 CAAAGTTTGGAGAGCTTGAACCTCACCATGACGACAGTTCAATGATTGTGTCAPCCFTGTG 788
|||||
Db 14 CRAAGTTTGGAGAGCTTGAACCTCACCATGACGACAGTTCAATGATTGTGTCATCTGTG 73
|||||
QY 789 TGGATGTGACTATTGTGATAGCATCAAAAGGTATCGGGGGCAACAGCTCTGAAACTTAT 848
|||||
Db 74 TGGATGTGACTATTGTGATAGCATCAAAAGGTATCGGGGGCAACAGCTCTGAAACTTAT 133
|||||
QY 849 TCGTCAACATGGTCCATAGAAGCATC 876
|||||
Db 134 TCGTCAACATGGTCCATAGAAGCATC 161
|||||

RESULT 18
AI834484/c 553 bp mRNA linear EST 02-FEB-2000
LOCUS 606068609.x1 606 - Ear tissue cDNA library from Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION AI834484
VERSION AI834484.1 GI:5468693
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 553)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 606068 row: G column: 09.
FEATURES
Source
Location/Qualifiers
1..553
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="606 - Ear tissue cDNA library from Schmidt
lab"
/tissue_type="mixed"
/dev_stage="ear length from 0.5 cm - 2.0 cm"
/lab_host="XLOLR (Stratagene)"
/note="Organ: immature ear; Vector: pBK-CMV; Site_1: EcoRI
; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt
lab"

BASE COUNT 140 a 112 c 106 g 195 t
ORIGIN

Query Match 9.1%; Score 134; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. NO. 3.3e-20;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 AGGGTCTCATAAGTTTCTCGTTAAAGATAATGTTTCAATGAAGATCGGGTGACAAAGG 1066
|||||
Db 553 AGGGTCTCATAAGTTTCTCGTTAAAGATAATGTTTCAATGAAGATCGGGTGACAAAGG 494
|||||
QY 1067 CCATAGAGAGATCAAAATCTCCAGAGATAATCGTCGCRAGGAAGACTCGAGTCCTTTT 1126
|||||
Db 493 CCATAGAGAGATCAAAATCTCCAGAGATAATCGTCGCRAGGAAGACTCGAGTCCTTTT 434
|||||
QY 1127 TCAAGCCAACTGCC 1140
|||||
Db 433 TCAAGCCAACTGCC 420
|||||

Best Local Similarity 100.0%; Pred. No. 3.8e-23;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 729 CAAAGTTTGGAGAGCTTGAACCTCACCATGACGACAGTTCAATGATTGTGTCAPCCFTGTG 788
|||||
Db 14 CRAAGTTTGGAGAGCTTGAACCTCACCATGACGACAGTTCAATGATTGTGTCATCTGTG 73
|||||
QY 789 TGGATGTGACTATTGTGATAGCATCAAAAGGTATCGGGGGCAACAGCTCTGAAACTTAT 848
|||||
Db 74 TGGATGTGACTATTGTGATAGCATCAAAAGGTATCGGGGGCAACAGCTCTGAAACTTAT 133
|||||
QY 849 TCGTCAACATGGTCCATAGAAGCATC 876
|||||
Db 134 TCGTCAACATGGTCCATAGAAGCATC 161
|||||

RESULT 18
AI834484/c 553 bp mRNA linear EST 02-FEB-2000
LOCUS 606068609.x1 606 - Ear tissue cDNA library from Schmidt lab Zea
DEFINITION mays cDNA, mRNA sequence.
ACCESSION AI834484
VERSION AI834484.1 GI:5468693
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 553)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 606068 row: G column: 09.
FEATURES
Source
Location/Qualifiers
1..553
/organism="Zea mays"
/cultivar="Ohio43"
/db_xref="taxon:4577"
/clone_lib="606 - Ear tissue cDNA library from Schmidt
lab"
/tissue_type="mixed"
/dev_stage="ear length from 0.5 cm - 2.0 cm"
/lab_host="XLOLR (Stratagene)"
/note="Organ: immature ear; Vector: pBK-CMV; Site_1: EcoRI
; Site_2: XhoI; Mixed ear tissue cDNA library from Schmidt
lab"

BASE COUNT 140 a 112 c 106 g 195 t
ORIGIN

Query Match 9.1%; Score 134; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. NO. 3.3e-20;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 AGGGTCTCATAAGTTTCTCGTTAAAGATAATGTTTCAATGAAGATCGGGTGACAAAGG 1066
|||||
Db 553 AGGGTCTCATAAGTTTCTCGTTAAAGATAATGTTTCAATGAAGATCGGGTGACAAAGG 494
|||||
QY 1067 CCATAGAGAGATCAAAATCTCCAGAGATAATCGTCGCRAGGAAGACTCGAGTCCTTTT 1126
|||||
Db 493 CCATAGAGAGATCAAAATCTCCAGAGATAATCGTCGCRAGGAAGACTCGAGTCCTTTT 434
|||||
QY 1127 TCAAGCCAACTGCC 1140
|||||
Db 433 TCAAGCCAACTGCC 420
|||||

Best Local Similarity 100.0%; Pred. No. 3.8e-23;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 729 CAAAGTTTGGAGAGCTTGAACCTCACCATGACGACAGTTCAATGATTGTGTCAPCCFTGTG 788
|||||
Db 14 CRAAGTTTGGAGAGCTTGAACCTCACCATGACGACAGTTCAATGATTGTGTCATCTGTG 73
|||||
QY 789 TGGATGTGACTATTGTGATAGCATCAAAAGGTATCGGGGGCAACAGCTCTGAAACTTAT 848
|||||
Db 74 TGGATGTGACTATTGTGATAGCATCAAAAGGTATCGGGGGCAACAGCTCTGAAACTTAT 133
|||||
QY 849 TCGTCAACATGGTCCATAGAAGCATC 876
|||||
Db 134 TCGTCAACATGGTCCATAGAAGCATC 161
|||||

RESULT 19
BF727781 232 bp mRNA linear EST 09-JAN-2001
LOCUS 1000052F12.x2 1000 - Unigene I from Maize Genome Project Zea mays
DEFINITION cDNA, mRNA sequence.
ACCESSION BF727781
VERSION BF727781.1 GI:12045642
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 232)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 1000052 row: F column: 12.
FEATURES
Source
Location/Qualifiers
1..232
/organism="Zea mays"
/db_xref="dbEST:707010C02.y1"
/db_xref="taxon:4577"
/clone_lib="1000 - Unigene I from Maize Genome Project"
/note="This library represents the unique ESTs found in
the first round of EST sequencing at Stanford University
for the maize genome project. Sequences are present from
libraries 486, 487, 496, 603, 605, 606, 614, 618, 660, 683
, 687, 707, and 945. Contigs were assembled using TIGR's
CAP program and a representative EST from each contig was
selected for the Unigene set. All singlets were also
selected."

BASE COUNT 64 a 42 c 57 g 69 t
ORIGIN

Query Match 7.0%; Score 103; DB 10; Length 232;
Best Local Similarity 100.0%; Pred. No. 3.2e-13;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 767 TCATTGATTTCGTCATCTCGTGTGATGTGACTATTGTGATAGCATCAAGGTATCGGG 826
|||||
Db 55 TCATTGATTTCGTCATCTCGTGTGATGTGACTATTGTGATAGCATCAAGGTATCGGG 114
|||||
QY 827 GGCAACAGCTCTGAAACTTATTCGTCACATCGGTCCTATAGA 869
|||||
Db 115 GGCAACAGCTCTGAAACTTATTCGTCACATCGGTCCTATAGA 157
|||||

RESULT 20
AI932215/c 105 bp mRNA linear EST 02-AUG-1999
LOCUS 618029H11.x1 618 - Inbred Tassel cDNA Library Zea mays cDNA, mRNA
DEFINITION sequence.
ACCESSION AI932215
VERSION AI932215.1 GI:5670952
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 105)
AUTHORS Walbot,V.

```

```

TITLE      Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL    University
COMMENT    Unpublished (1999)
           Contact: Walbot V
           Department of Biological Sciences
           Stanford University
           855 California Ave, Palo Alto, CA 94304, USA
           Tel: 650 723 2227
           Fax: 650 725 8221
           Email: walbot@stanford.edu
           Plate: 618029 row: H column: 11.
           Location/Qualifiers
             1..105
               /organism="Zea mays"
               /cultivar="Ohio43"
               /db_xref="taxon:4577"
               /clone_lib="618 - Inbred Tassel cDNA Library"
               /tissue_type="tassel"
               /dev_stage="tassel"
               /lab_host="XL0LR"
               /note="Organ: tassel; Vector: pAD-GAL4-2.1 (Hybrizap);
               Inbred tassel library from Schmidt lab"
BASE COUNT 36 a 27 c 21 g 21 t
ORIGIN
Query Match 4.6%; Score 68; DB 9; Length 105;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1309 CTCCTGTTTAACTCAGACCTTGGTGAAGTTTGCCCATGTTTCAAGCTGGGGTAAGT 1368
Db 68 CTCCTGTTTAACTCAGACCTTGGTGAAGTTTGCCCATGTTTCAAGCTGGGGTAAGT 9

QY 1369 TAGTTGTG 1376
Db 8 TAGTTGTG 1

RESULT 21
AW147048
LOCUS      AW147048 126 bp mRNA linear EST 03-NOV-1999
DEFINITION      707010C02.y1 707 - Mixed adult tissues from Walbot lab (SK) Zea
mays cDNA, mRNA sequence.
ACCESSION      AW147048
VERSION        AW147048.1 GI:6194944
KEYWORDS       EST.
SOURCE         Zea mays.
ORGANISM       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
Walbot.V.
Maize ESTs from various cDNA libraries sequenced at Stanford

JOURNAL      Unpublished (1999)
COMMENT      Contact: Walbot V
             Department of Biological Sciences
             Stanford University
             855 California Ave, Palo Alto, CA 94304, USA
             Tel: 650 723 2227
             Fax: 650 725 8221
             Email: walbot@stanford.edu
             Plate: 707010 row: C column: 02.
             Location/Qualifiers
               1..126
                 /organism="Zea mays"
                 /cultivar="W23"
                 /db_xref="taxon:4577"
                 /clone_lib="707 - Mixed adult tissues from Walbot lab (SK)
                 )"
                 /tissue_type="tassel, kernel, silk, husk, root, leaf"
                 /dev_stage="adult"

FEATURES     source
             Location/Qualifiers
               1..126
                 /organism="Zea mays"
                 /cultivar="W23"
                 /db_xref="taxon:4577"
                 /clone_lib="707 - Mixed adult tissues from Walbot lab (SK)
                 )"
                 /tissue_type="tassel, kernel, silk, husk, root, leaf"
                 /dev_stage="adult"

```

```

/lab_host="DH10B"
/note="Organ: tassel, kernel, silk, husk, root, leaf;
Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
differentiated maize tissues from an active Mutator
plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
husk, root, leaf). Unidirectionally cloned."
BASE COUNT 29 a 31 c 26 g 40 t
ORIGIN
Query Match 4.6%; Score 68; DB 9; Length 126;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 809 GCATCAAGGTATCGGGGGCAACAGCTCTGAAACTTATTGTCACATGGTCCATAG 868
Db 1 GCATCAAGGTATCGGGGGCAACAGCTCTGAAACTTATTGTCACATGGTCCATAG 60

QY 869 AAAGCATC 876
Db 61 AAAGCATC 68

RESULT 22
AW147048/c
LOCUS      AW147048 126 bp mRNA linear EST 03-NOV-1999
DEFINITION      707010C02.y1 707 - Mixed adult tissues from Walbot lab (SK) Zea
mays cDNA, mRNA sequence.
ACCESSION      AW147048
VERSION        AW147048.1 GI:6194944
KEYWORDS       EST.
SOURCE         Zea mays.
ORGANISM       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
Walbot.V.
Maize ESTs from various cDNA libraries sequenced at Stanford

JOURNAL      Unpublished (1999)
COMMENT      Contact: Walbot V
             Department of Biological Sciences
             Stanford University
             855 California Ave, Palo Alto, CA 94304, USA
             Tel: 650 723 2227
             Fax: 650 725 8221
             Email: walbot@stanford.edu
             Plate: 707010 row: C column: 02.
             Location/Qualifiers
               1..126
                 /organism="Zea mays"
                 /cultivar="W23"
                 /db_xref="taxon:4577"
                 /clone_lib="707 - Mixed adult tissues from Walbot lab (SK)
                 )"
                 /tissue_type="tassel, kernel, silk, husk, root, leaf"
                 /dev_stage="adult"
                 /lab_host="DH10B"
                 /note="Organ: tassel, kernel, silk, husk, root, leaf;
                 Vector: pGAD10; Site_1: EcoRI; cDNA library from fully
                 differentiated maize tissues from an active Mutator
                 plant. Tissue ratio is 4/2/1/1/1 (tassel, kernel, silk,
                 husk, root, leaf). Unidirectionally cloned."
BASE COUNT 29 a 31 c 26 g 40 t
ORIGIN
Query Match 4.4%; Score 65; DB 9; Length 126;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1184 CAAGCAAGCGCCTCGCAACAGAAACAAAGCTGCTGGTGAAGAAAGAAATATCTTGA 1243
Db 126 CAAGCAAGCGCCTCGCAACAGAAACAAAGCTGCTGGTGAAGAAAGAAATATCTTGA 67

```

```

QY 1244 TGCTT 1248
      |||||
Db 66 TGCTT 62

RESULT 23
AI947478/c
LOCUS
DEFINITION 614047B01.xl 614 - root cDNA library from Walbot Lab Zea mays cDNA,
      mRNA sequence.
ACCESSION AI947478
VERSION AI947478.1 GI:57396683
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Zea mays.
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
      clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 225)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
      University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
      Department of Biological Sciences
      Stanford University
      855 California Ave, Palo Alto, CA 94304, USA
      Tel: 650 723 2227
      Fax: 650 725 8221
      Email: walbot@stanford.edu
      Plate: 614047 row: B column: 01.
      Location/Qualifiers
      1..225
      /organism="Zea mays"
      /cultivar="W23"
      /db_xref="taxon:4577"
      /clone_lib="614 - root cDNA library from Walbot Lab"
      /tissue_type="root"
      /dev_stage="3-4 days old"
      /lab_host="XL0LR"
      /note="Organ: root; Vector: pBluescriptII SK+; Site:1:
      EcoRI; Site_2: XhoI; 3-4 days old root tissue from Walbot
      Lab (LM)"
BASE COUNT 66 a 54 c 40 g 65 t
ORIGIN

      Query Match 3.9%; Score 57; DB 9; Length 225;
      Best Local Similarity 100.0%; Pred. No. 0.0024;
      Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1197 TCGCACACAGAAACAAAGCGCTGGTGGAAAGAGAAATAATCTGGATGCTTGATGT 1253
      |||||
Db 225 TCGCAACAGAAACAAAGCGCTGGTGGAAAGAGAAATAATCTGGATGCTTGATGT 169

RESULT 24
D43467
LOCUS
DEFINITION D43467 Rice callus cDNA (H.Uchimiya) Oryza sativa cDNA clone SS447,
      mRNA sequence.
ACCESSION D43467
VERSION D43467.1 GI:3107727
KEYWORDS EST.
SOURCE Oryza sativa.
ORGANISM Oryza sativa
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
      Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 253)
AUTHORS Uchimiya,H.
TITLE On nucleotide sequence of Oryza sativa
JOURNAL Unpublished (1994)

Contact: Hirofumi Uchimiya
Institute of Mol. & Cell. Bioscience, Department of Cellular
Function
The University of Tokyo
1-1-1 Yayoi, Bunkyo-ku, Tokyo 113, Japan
Tel: 03-3812-2111(ex.7844)
Fax: 03-3812-2910
Email: huchimiya@tansei.cc.u-tokyo.ac.jp
PROJECT= Uchimiya.
      Location/Qualifiers
      1..253
      /organism="Oryza sativa"
      /db_xref="taxon:4530"
      /clone="SS447"
      /clone_lib="Rice callus cDNA (H.Uchimiya)"
      /tissue_type="callus"
      74 a 42 c 64 g 74 t 1 others
BASE COUNT 72 a 42 c 64 g 74 t 1 others
ORIGIN

      Query Match 3.0%; Score 44; DB 10; Length 253;
      Best Local Similarity 100.0%; Pred. No. 1.3;
      Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 520 TGTAAACGACTATTAAAGACTTATGGGGTTCCTGCTGTAGAGGC 563
      |||||
Db 25 TGTAAACGACTATTAAAGACTTATGGGGTTCCTGCTGTAGAGGC 68

RESULT 25
AV915644
LOCUS
DEFINITION AV915644 K. Sato unpublished cDNA library, cv. Haruna NiJo
      germination shoots Hordeum vulgare subsp. vulgare cDNA clone
      bags15h10 5', mRNA sequence.
ACCESSION AV915644
VERSION AV915644.1 GI:18211421
KEYWORDS EST.
ORGANISM Hordeum vulgare subsp. vulgare.
      Hordeum vulgare subsp. vulgare.
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
      Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 454)
AUTHORS Sato,K., Saisho,D. and Takeda,K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
      Center For Genetic Resource Information
      National Institute of Genetics
      1111 Yata, Mishima, Shizuoka 411-8540, Japan
      Tel: 81-559-81-6856
      Fax: 81-559-81-6855
      Email: tshiniegenes.nig.ac.jp.
      Location/Qualifiers
      1..454
      /organism="Hordeum vulgare subsp. vulgare"
      /cultivar="Haruna NiJo"
      /db_xref="taxon:112509"
      /clone="bags15h10"
      /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
      NiJo germination shoots"
      /tissue_type="shoots"
      /dev_stage="germination"
BASE COUNT 140 a 90 c 125 g 99 t
ORIGIN

      Query Match 2.6%; Score 38; DB 9; Length 454;
      Best Local Similarity 100.0%; Pred. No. 15;
      Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATCGCGGTGACGCCAGCATGAGCATCTACCAGTTCCT 224
      |||||
Db 17 ATCGCGGTGACGCCAGCATGAGCATCTACCAGTTCCT 54

```

```

RESULT 26
AV913663
LOCUS
DEFINITION
    AV913663 K. sato unpublished cDNA library, linear EST 18-JAN-2002
    germination shoots Hordeum vulgare subsp. vulgare cDNA clone
    bags22n17 5', mRNA sequence.
ACCESSION
    AV913663
VERSION
    1
KEYWORDS
    EST.
ORGANISM
    Hordeum vulgare subsp. vulgare.
    Hordeum vulgare subsp. vulgare
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
    ; Triticeae; Hordeum.
    1 (bases 1 to 648)
REFERENCE
    Sato,K., Saisho,D. and Takeda,K.
    Barley EST sequencing project in NIG and Okayama Univ
    Unpublished (2002)
    Contact: Tadasi Shin-1
    Center For Genetic Resource Information
    National Institute of Genetics
    1111 Yata, Mishima, Shizuoka 411-8540, Japan
    Tel: 81-559-81-6856
    Fax: 81-559-81-8855
    Email: tshin@genes.nig.ac.jp.
    Location/Qualifiers
    1..648
    /organism="Hordeum vulgare subsp. vulgare"
    /cultivar="Haruna Nijo"
    /db_xref="taxon:112509"
    /clone="bags22n17"
    /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
    Nijo germination shoots"
    /tissue_type="shoots"
    /dev_stage="germination"
    174 a 158 c 131 t

BASE COUNT      174 a 158 c 131 t
ORIGIN
    Query Match      2.6%; Score 38; DB 9; Length 648;
    Best Local Similarity 100.0%; Pred. No. 11;
    Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 187 ATGCGCGTCGACGCCAGCATGAGCATCTACCACTTCT 224
      |||||||||||||||||||||||||||||||||||
Db 172 ATGCGCGTCGACGCCAGCATGAGCATCTACCACTTCT 209

RESULT 27
BF098320
LOCUS
DEFINITION
    BF098320 tomato nutrient deficient roots Lycopersicon esculentum
    cDNA clone cLEW26K15 5' sequence similar to 5 nuclease xFEN1a
    (Xenopus laevis) GP141063581gb|AAD02814.1|AF065397 flap
    endonuclease 1 (Xenopus laevis) mRNA sequence.
ACCESSION
    BF098320
VERSION
    1
KEYWORDS
    EST.
SOURCE
    tomato.
    Lycopersicon esculentum
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
    Lycopersicon.
    1 (bases 1 to 539)
REFERENCE
    van der Hoeven,R.S., Garvin,D.F., Matern,A.L., Holt,I.E., Liang,F.,
    Upton,J., Hansen,T.S., Ronning,C.M., Craven,M.B., Bowman,C.L.,
    Niernan,W., Fraser,C.M., Venter,J.C., Martin,G.B., Giovannoni,J.J.
    and Tanksley,S.D.
    Generation of ESTs from tomato nutrient-deficient roots
    Unpublished (1999)
    Contact: CUGI

TITLE
    Journal
    COMMENT
    BF098320

```

```

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
    1..539
    /organism="Lycopersicon esculentum"
    /cultivar="74492"
    /db_xref="taxon:4081"
    /clone="cLEW26K15"
    /clone_lib="tomato nutrient deficient roots"
    /tissue_type="roots"
    /dev_stage="5-6 weeks old"
    /lab_host="SOLR"
    /note="vector: pBluescriptSKmCuaadapt; site1: 5' EcoRI;
    Site 2: 3' XhoI; Roots were harvested from plants grown
    under the following deficiencies/stresses: 10 mM Al, Zn,
    P, K, Fe.N. mRNA was isolated from individual treatments.
    Proportional aliquots of mRNA of each treatment were mixed
    and used for library construction."
    BASE COUNT      165 a 97 c 143 g 134 t
    ORIGIN
    Query Match      2.4%; Score 35; DB 10; Length 539;
    Best Local Similarity 100.0%; Pred. No. 57;
    Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 GGAATCAAGCCAGTTATGTTTTGATGGCAAGCC 365
      |||||||||||||||||||||||||||||||
Db 273 GGAATCAAGCCAGTTATGTTTTGATGGCAAGCC 307

RESULT 28
AT780966
LOCUS
DEFINITION
    AT780966 tomato susceptible, Cornell Lycopersicon esculentum cDNA
    clone cLES13P2, mRNA sequence.
ACCESSION
    AT780966
VERSION
    1
KEYWORDS
    EST.
SOURCE
    tomato.
    Lycopersicon esculentum
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
    Lycopersicon.
    1 (bases 1 to 592)
REFERENCE
    D' Ascenzo,M., He,X., Lyman,J., Matern,A.L., Vision,T., Holt,I.E.,
    Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
    ,C.L., Niernan,W., Fraser,C.M., Venter,J.C., Tanksley,S.D.,
    Giovannoni,J.J. and Martin,G.B.
    Generation of ESTs from Pseudomonas susceptible tomato
    Unpublished (1999)
    Contact: CUGI
    Clemson University Genomics Institute
    Clemson University
    100 Jordan Hall, Clemson, SC 29634, USA
    Email: http://www.genome.clemson.edu/orders/index.html
    5 prime sequence.
    Location/Qualifiers
    1..592
    /organism="Lycopersicon esculentum"
    /cultivar="Ril-13 (Rio Grande x Money Maker)"
    /db_xref="taxon:4081"
    /clone="cLES13P2"
    /clone_lib="tomato susceptible, Cornell"
    /tissue_type="leaf"
    /dev_stage="4-week old"
    /lab_host="SOLR"
    /note="vector: pBluescript SK(-); Site1: EcoRI; Site2:
    XhoI; cLES - Tomato Pseudomonas Susceptible EST Library.
    Directionally cloned cDNAs inserted into pBluescript SK(-
    ) at 5' end with EcoRI and 3' end with XhoI site"

```

```

BASE COUNT      190 a      109 c      149 g      144 t
ORIGIN
Query Match      2.4%; Score 35; DB 9; Length 592;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 331 GGAATCAAGCCAGTTATGTTTGTGATGGCAAGCC 365
|||||
Db 306.GGAATCAAGCCAGTTATGTTTGTGATGGCAAGCC 340

RESULT 29
AI271296      143 bp      mRNA      linear      EST 21-DEC-1998
LOCUS      qw69g01.x1 NCI_CGAP_Ov33 Homo sapiens cDNA clone IMAGE:1996272 3',
DEFINITION      mRNA sequence.
ACCESSION      AI271296
VERSION      AI271296.1 GI:3890463
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 143)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: David B. Krizman, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 239 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 135.
Location/Qualifiers
1..143
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1996272"
/sex="female"
/tissue_type="borderline ovarian carcinoma"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: ovary; Vector: pAMP1; mRNA made from
borderline ovarian carcinoma, cDNA made by oligo-dT
priming. Directionally cloned. Size-selected on agarose
gel, average insert size 500 bp. Primary library,
non-amplified."
BASE COUNT      52 a      34 c      32 g      24 t      1 others
ORIGIN

Query Match      2.1%; Score 31; DB 9; Length 143;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1448 TGCCAGTAAAAAATAAAAAAAAAAAAAAAAAA 1478
|||||
Db 110 TGCCAGTAAAAAATAAAAAAAAAAAAAAAAAA 140

RESULT 30
AI308412      152 bp      mRNA      linear      EST 08-APR-1999
LOCUS      tb43e12.x1 NCI_CGAP_Br16 Homo sapiens cDNA clone IMAGE:2057134 3',
DEFINITION      mRNA sequence.
ACCESSION      AI308412

```

```

AI308412.1 GI:4003047
EST.
human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 152)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dennis Sgroi, M.D., Kristina Cole, M.D., Ph.D.
student, Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 234 Std Error: 0.00
Seq primer: -400P from Gibco.
Location/Qualifiers
1..152
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2057134"
/sex="female"
/tissue_type="lobullar carcinoma in situ"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: breast; Vector: pAMP1; mRNA made from breast
carcinoma tissue, cDNA made by oligo-dT priming.
Directionally cloned. Size-selected on agarose gel,
average insert size 400 bp. Primary library,
non-amplified."
BASE COUNT      53 a      37 c      35 g      27 t
ORIGIN

Query Match      2.1%; Score 31; DB 9; Length 152;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1448 TGCCAGTAAAAAATAAAAAAAAAAAAAAAAAA 1478
|||||
Db 110 TGCCAGTAAAAAATAAAAAAAAAAAAAAAAAA 140

RESULT 31
AI305700      157 bp      mRNA      linear      EST 03-FEB-1999
LOCUS      qw47e05.x1 NCI_CGAP_Br13 Homo sapiens cDNA clone IMAGE:1994240 3',
DEFINITION      mRNA sequence.
ACCESSION      AI305700
VERSION      AI305700.1 GI:3990591
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 157)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: David B. Krizman, Ph.D.
DNA Sequencing by: I.M.A.G.E. Consortium, LLNL
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be

```


found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.lnl.gov/brp/image/image.html
 Insert Length: 237 Std Error: 0.00
 Seq primer: -40UP from Gibco.

FEATURES

source

1. .157
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1994240"
 /clone_lib="NCI_CGAP_Brl3"
 /sex="female"
 /tissue_type="breast carcinoma in situ"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: breast; Vector: pAMP1; mRNA made from breast carcinoma. cDNA made by oligo-dT priming. Directionally cloned. Size selected on agarose gel, average insert size 500 bp. Primary library, non-amplified."
 58 a 37 c 35 g 27 t

BASE COUNT

ORIGIN

Query Match 2.1%; Score 31; DB 9; Length 157;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1448 TCCAGTAAAAA 1478
 Db 110 TCCAGTAAAAA 140

RESULT 32

AI250739

LOCUS

gx23d02.x1 NCI_CGAP_Ov34 163 bp mRNA linear. EST 05-NOV-1998
 mRNA sequence.

ACCESSION

AI250739

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 unknown library type
 Seq primer: -40UP from Gibco.

FEATURES

source

1. .163
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2002179"
 /clone_lib="NCI_CGAP_Ov34"
 /sex="female"
 /tissue_type="borderline ovarian carcinoma"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Organ: ovary; Vector: pAMP1; mRNA made from borderline ovarian carcinoma, cDNA made by oligo-dT priming. Directionally cloned. Size-selected on agarose gel, average insert size 500 bp. Primary library, non-amplified."
 55 a 40 c 38 g 30 t

BASE COUNT

ORIGIN

Query Match 2.1%; Score 31; DB 9; Length 163;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1448 TCCAGTAAAAA 1478
 Db 121 TCCAGTAAAAA 151

RESULT 33

AI266300

LOCUS

qp69c03.x1 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
 IMAGE:1928260 3', mRNA sequence.

ACCESSION

AI266300

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1164 Std Error: 0.00
 Seq primer: -40UP from Gibco.

FEATURES

source

1. .179
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1928260"
 /clone_lib="Soares_fetal_lung_NbHL19W"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: lung; Vector: p7T3D (Pharmacia) with a modified polylinker; Site: 1; Not I; Site: 2; Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTAGCAATCTGAAGTGGAGCGCGCAATTTTTTTTTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NbHL19W."
 62 a 42 c 41 g 34 t

BASE COUNT

ORIGIN

Query Match 2.1%; Score 31; DB 9; Length 179;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1448 TCCAGTAAAAA 1478

Db 127 TCCAGTAAAAA 157

RESULT 34

AI825335/c

LOCUS

wb17b09.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2305913 3',
 mRNA sequence.

DEFINITION

ACCESSION

AI825335

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1 (bases 1 to 253)

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.


```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 2166 Std Error: 0.00
Seq primer: -40m13 fwd. Et from AmerSham
High quality sequence stop: 304.
Location/Qualifiers
1. .326
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1676813"
/clone_lib="Soares_Total_fetus_Nb2HF8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/Note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TCTTACCAATCTGAAGTCGGAGCGCGCTTAATTTTATTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaído. "
BASE COUNT 128 a 44 c 45 g 109 t
ORIGIN
Query Match 2.1%; Score 31; DB 9; Length 326;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1448 TGCAGTAAAAA...AAAAAAAAA 1478
|||||
Db 32 TGCAGTAAAAA...AAAAAAAAA 2

RESULT 37
AI394973/c
LOCUS AI394973 366 bp mRNA linear EST 04-FEB-1999
DEFINITION MA001960.c8f Soares normalized 88w Schistosoma mansoni cDNA 3',
mRNA sequence.
VERSION AI394973.1 GI:4224520
KEYWORDS EST.
SOURCE Schistosoma mansoni.
ORGANISM Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.
REFERENCE
AUTHORS Bailey,J.A., Bentley,K., Blanton,R.E., Soares,M.B. and Chakravarti
,A.
TITLE Expressed sequence tags from a normalized adult Schistosoma mansoni
library
JOURNAL Unpublished (1999)
COMMENT Contact: Chakravarti A
Department of Genetics
Case Western Reserve University
2109 Adelbert Rd, Cleveland, OH 44106, USA
Tel: 216-368-5847
Fax: 216-368-5857
Email: axc39@po.cwru.edu
Additional data regarding this EST may be found at
http://genome.cwru.edu/schistosoma/est/S8West.html
Seq primer: M13 forward.
Location/Qualifiers
1. .366

```

```

/organism="Schistosoma mansoni"
/strain="Puerto Rican"
/db_xref="taxon:6183"
/clone_lib="Soares normalized 88w"
/sex="mixed"
/dev_stage="8 week old adult worms"
/lab_host="Mus musculus"
/Note="vector: pT7T3-pac; Site_1: NotI; Site_2: EcoRI;
Note: normalization has resulted in an increased frequency
of the rare contaminating mouse (host) sequences (~10%).
For details of library construction see: Bonatido MF,
Lennon G, Soares MB. Normalization and Subtraction: Two
Approaches to Facilitate Gene Discovery. 1996. Genome
Research 6:791-806"
BASE COUNT 99 a 82 c 68 g 117 t
ORIGIN
Query Match 2.1%; Score 31; DB 9; Length 366;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1448 TGCAGTAAAAA...AAAAAAAAA 1478
|||||
Db 32 TGCAGTAAAAA...AAAAAAAAA 2

RESULT 38
AI165690
LOCUS AI165690 408 bp mRNA linear EST 06-DEC-2000
DEFINITION AU165690 Rice panicle at flowering stage Oryza sativa cDNA clone
E4061, mRNA sequence.
ACCESSION AU165690
VERSION AU165690.1 GI:11565054
KEYWORDS EST
SOURCE Oryza sativa.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE
AUTHORS Sasaki,T. and Yamamoto,K.
TITLE Rice cDNA from panicle at flowering stage (2000)
JOURNAL Unpublished (2000)
COMMENT Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT = 'RGP'
E4061_4Z. Location/Qualifiers
1. .408
/organism="Oryza sativa"
/strain="Nipponbare"
/db_xref="taxon:4530"
/clone="E4061"
/clone_lib="Rice panicle at flowering stage"
/dev_stage="flowering stage"
/Note="Organ: panicle; Rice cDNA from panicle at flowering
stage"
BASE COUNT 119 a 78 c 93 g 116 t 2 others
ORIGIN
Query Match 2.1%; Score 31; DB 9; Length 408;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1448 TGCAGTAAAAA...AAAAAAAAA 1478
|||||
Db 366 TGCAGTAAAAA...AAAAAAAAA 396

```

```

RESULT 39
AI017545/c      444 bp      mRNA      linear      EST 27-AUG-1998
LOCUS           Ou35c07.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
DEFINITION      IMAGE:1628268 3', mRNA sequence.
ACCESSION      AI017545
VERSION         AI017545.1      GI:3231881
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 444)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
COMMENT        Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-remail.nih.gov
                This clone is available royalty-free through LLNL; contact the
                IMAGE Consortium (infoimage.llnl.gov) for further information.
                Insert Length: 776      Std Error: 0.00
                Seq primer: -40m13 fwd. ET from Amersham
                High quality sequence stop: 414.
FEATURES       Location/Qualifiers
                1..444
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:1628268"
                /clone_lib="Soares_NFL_T_GBC_S1"
                /lab_host="DH10B"
                /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                Equal amounts of plasmid DNA from three normalized
                libraries (fetal lung NBHL19W, testis NHT, and B-cell
                NCLCGAP_GCB1) were mixed and ss circles were made in
                vitro. Following HAP purification, this DNA was used as
                tracer in a subtractive hybridization reaction. The driver
                was PCR-amplified cDNAs from pools of 5,000 clones made
                from the same 3 libraries. The pools consisted of
                I.M.A.G.E. clones 297480-302087, 682632-687239,
                726408-728711, and 729096-731399. Subtraction by Bento
                Soares and M. Fatima Bonaldo."
                BASE COUNT      135 a      68 c      61 g      180 t
ORIGIN
Query Match      2.1%; Score 31; DB 9; Length 444;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1448 TGCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
|||||
Db 31 TGCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1

RESULT 40
AW019727/c      479 bp      mRNA      linear      EST 07-JUN-2001
LOCUS           fd56d01.x1 Zebrafish Washu MPIMG EST Danio rerio cDNA clone
DEFINITION      IMAGE:373729 3', mRNA sequence.
ACCESSION      AW019727
VERSION         AW019727.1      GI:5873257
KEYWORDS       EST.
SOURCE         zebrafish.
ORGANISM       Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
1 (bases 1 to 479)
Clark.M., Johnson,S.L., Lehrach.H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wyllie,T., Underwood
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,

```

Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
Washu Zebrafish EST Project 1998
Unpublished (1998)
Other_ESTs: fd56d01.y1
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrfish@watson.wustl.edu
CDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone Distribution: Genome Systems, St. Louis,
Missouri (web address: www.genomesystems.com) (email contact:
info@genomesystems.com) and Research Genetics, Huntsville, Alabama
(web address: www.resgen.com) (email contact: info@resgen.com) and
RessourcenzentrumPrimarDatenbank, Berlin, Germany (web address:
www.rzpd.de)

Seq primer: T7 ET from Amersham
High quality sequence stop: 438.

FEATURES
source ,

1..479
Location/Qualifiers
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="IMAGE:373729"
/clone_lib="Zebrafish Washu MPIMG EST"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield
stage embryos"
/lab_host="XLI-blue MRF"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
strand cDNA was primed with a Not I - oligo(dT)15 primer
[5'pGACTAGTCTAGATCGAGCGCGCCCTTTTCTTTT3'];
double-stranded cDNA was ligated to Sal I adaptors (BRL),
digested with Not I and cloned into the Not I and Sal I
sites of the pSPORT1 vector (BRL). Library was constructed
by Matthew Clark (Lehrach lab; ICRF, London and Max Planck
Institut fuer Molekulare Genetik, Berlin). cDNAs for EST
analysis were selected following oligonucleotide
hybridization fingerprinting of arrayed clones from
embryonic shield stage (5.6 h) libraries. Fingerprint
data were used to computationally cluster cDNAs, and a
single cDNA from each cluster was chosen for sequencing.
In some cases multiple members of the same cluster were
sequenced to assess clustering parameters or single clones
were sequenced additional times to assess quality
control."

BASE COUNT 185 a 80 c 57 g 157 t
ORIGIN

Query Match 2.1%; Score 31; DB 9; Length 479;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1448 TGCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 1478
|||||
Db 39 TGCAGTAAATAAAAAAAAAAAAAAAAAAAAAA 9

RESULT 41

AA454191
LOCUS zx48b10.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:795451
DEFINITION 5', mRNA sequence.
ACCESSION AA454191
VERSION AA454191.1 GI:2167860
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
              1. (bases 1 to 373)
              Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
              Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
              Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie,
              T., Waterston, R., and Wilson, R.
              WashU-Merck EST Project 1997
              Unpublished (1997)
              Contact: Wilson RK
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              This clone is available royalty-free through LLNL; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              Seq primer: -28m13 rev2 ET from Amersham
              High quality sequence stop: 500.

FEATURES
source
1..573
   /organism="Homo sapiens"
   /db_xref="GDB:6038625"
   /db_xref="taxon:9606"
   /clone_lib="IMAGE:795451"
   /clone_lib="Soares_testis_NHT"
   /sex="male"
   /lab_host="DH10B"
   /note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTTCACCAATCTGAAGTGGGAGCGGCCGCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      230 a 109 c 97 g 137 t
ORIGIN
Query Match      2.1%; Score 31; DB 9; Length 573;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1448 TGCACGTAAAAA...AAAAAAAAAAAAA 1478
|||||
Db 504 TGCACGTAAAAA...AAAAAAAAAAAAA 534

RESULT 42
AW071753/c
LOCUS      640 bp mRNA linear EST 08-MAR-2000
DEFINITION w555c02.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2501090 3'
similar to TR:064850 064850 TID16.16 PROTEIN. ;, mRNA sequence.
ACCESSION  AW071753
VERSION     AW071753.1 GI:6026839
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 640)
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/PTGAP), Tumor Gene Index
Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.

```

```

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1211 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 334.

FEATURES
Location/Qualifiers
1..640
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone="IMAGE:2501090"
   /clone_lib="NCI_CGAP_Brn25"
   /tissue_type="anaplastic oligodendroglioma"
   /lab_host="DH10B"
   /note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTTCACCAATCTGAAGTGGGAGCGGCCGCAATTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT      183 a 142 c 127 g 187 t
ORIGIN
Query Match      2.1%; Score 31; DB 9; Length 640;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1448 TGCACGTAAAAA...AAAAAAAAAAAAA 1478
|||||
Db 34 TGCACGTAAAAA...AAAAAAAAAAAAA 4

RESULT 43
BM078557/c
LOCUS      681 bp mRNA linear EST 14-NOV-2001
DEFINITION MEST121-C06.T3 ISUM4-TN Zea mays cDNA clone MEST121-C06 3', mRNA
sequence.
ACCESSION  BM078557
VERSION     BM078557.1 GI:16925489
KEYWORDS   EST.
SOURCE     Zea mays.
ORGANISM   Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 681)
Olu, F., Cui, F., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
Expressed Sequence Tags from B73 Maize Seedlings and Silks
Unpublished (2001)
Contact: Patrick S. Schnable
Schnable Laboratory
Iowa State University
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-2299
Email: schnable@iastate.edu
Individual basecall and confidence value were assigned using the
Phred software,
(<http://depts.washington.edu/ventures/collabtr/direct/index.htm>#b
rt). Overall sequence quality assessment and vector trimming were
conducted using the Lucy software (<http://www.tigr.org/softlab/>).
Lucy parameters were set to ensure an overall trimmed quality of
97.5% or better without any vector fragments in the chosen
high-quality region of each sequence. Low-quality bases between the
poly-T and the high-quality region were replaced with N's to serve
as spacers.
PCR Primers
FORWARD: Forward PCR primer sequence, primer T7-1 (AA TAC GAC TCA

```


reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT	34 a	26 c	29 g	48 t
ORIGIN				
Query Match	2.0%; Score 30; DB 9; Length 137;			
Best Local Similarity	100.0%; Pred. No. 2.4e+03;			
Matches 30; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;
QY 1449	GCCAGTAAAAA 1478			
Db 32	GCCAGTAAAAA 3			

Search completed: November 5, 2002, 16:22:46
Job time : 1293.61 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 16:22:54 ; Search time 25 seconds

(without alignments)
370.292 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 379

Sequence: 1 MGKGLTKLLADNAPKAMKE.....SDKTSKAAANKTKRAGGKKK 379

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/6C_US_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	379	100.0	379	4	US-09-426-557-2
2	379	100.0	379	4	US-09-426-557-6
3	321	84.7	379	4	US-09-426-557-4
4	278	73.4	379	4	US-09-426-557-8
5	13	3.4	377	2	US-08-455-968E-3
6	13	3.4	378	2	US-08-455-968E-10
7	13	3.4	378	2	US-08-823-516-138
8	13	3.4	380	2	US-08-455-968E-1
9	13	3.4	380	2	US-08-823-516-137
10	13	3.4	380	4	US-09-426-557-10
11	12	3.2	340	2	US-08-757-653-176
12	12	3.2	340	2	US-08-823-516-79
13	12	3.2	340	2	US-08-823-516-136
14	12	3.2	340	3	US-08-759-038-115
15	12	3.2	340	3	US-08-758-314-115
16	9	2.4	16	2	US-08-455-968E-16
17	9	2.4	382	2	US-08-455-968E-5
18	9	2.4	382	2	US-08-823-516-139
19	8	2.1	343	4	US-09-146-319-2
20	8	2.1	343	4	US-09-175-973-2
21	7	1.8	25	2	US-08-455-968E-20
22	7	1.8	117	4	US-08-858-207A-453
23	7	1.8	179	6	5463025-4
24	7	1.8	364	4	US-09-077-675A-16
25	7	1.8	464	6	5463025-1
26	7	1.8	488	1	US-08-115-365-2
27	7	1.8	488	1	US-08-586-897-2

28	7	1.8	513	1	US-08-390-162-6	Sequence 6, Appli
29	7	1.8	513	1	US-08-685-945B-6	Sequence 6, Appli
30	7	1.8	555	2	US-08-588-258B-24	Sequence 24, Appli
31	7	1.8	555	3	US-08-460-505-24	Sequence 24, Appli
32	7	1.8	555	5	PCT-US96-08295-24	Sequence 24, Appli
33	7	1.8	638	2	US-08-557-122A-38	Sequence 38, Appli
34	7	1.8	638	4	US-09-262-666-38	Sequence 38, Appli
35	7	1.8	664	6	5240838-5	Patent No. 5240838
36	7	1.8	685	5	PCT-US91-09784-4	Sequence 4, Appli
37	7	1.8	973	2	US-08-683-262B-75	Sequence 75, Appli
38	7	1.8	973	4	US-09-361-707-75	Sequence 75, Appli
39	7	1.8	1121	4	US-09-171-461-28	Sequence 28, Appli
40	7	1.8	3969	4	US-08-061-376-5	Sequence 5, Appli
41	6	1.6	9	5	PCT-US95-04121-35	Sequence 35, Appli
42	6	1.6	14	2	US-08-591-438-12	Sequence 12, Appli
43	6	1.6	14	6	5405952-5	Patent No. 5405952
44	6	1.6	16	2	US-08-712-212-3	Sequence 3, Appli
45	6	1.6	16	2	US-08-712-212-12	Sequence 12, Appli

ALIGNMENTS

RESULT 1

US-09-426-557-2
Sequence 2, Application US/09426557
Patent No. 6232527

APPLICANT: MahaJan, Pramod B.
TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
TITLE OF INVENTION: Thereof
FILE REFERENCE: 0961
CURRENT APPLICATION NUMBER: US/09/426,557
CURRENT FILING DATE: 1999-10-22
EARLIER APPLICATION NUMBER: 60/112,332
EARLIER FILING DATE: 1998-12-15
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 379
TYPE: PRT
ORGANISM: Zea mays
US-09-426-557-2

Query Match 100.0%; Score 379; DB 4; Length 379;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MGKGLTKLLADNAPKAMKEOKFFSYFGKTAVDASMSIYQFLIVGRTGMEITLNEAGE	50
Db	1	MGKGLTKLLADNAPKAMKEOKFFSYFGKTAVDASMSIYQFLIVGRTGMEITLNEAGE	60
QY	61	VTSHLQGMFNRTIRLLEAGIRPVVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVG	120
Db	61	VTSHLQGMFNRTIRLLEAGIRPVVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVG	120
QY	121	KDAIEKLSKRTVKVTRQNHNECKRLRLRMGVVVEAPSEAECAALCINDKVFVASED	180
Db	121	KDAIEKLSKRTVKVTRQNHNECKRLRLRMGVVVEAPSEAECAALCINDKVFVASED	180
QY	181	MDSLTFGAPRFLRHLMDSSKKIPVMEFDVAKVLEELTMDQFIDLCILCGDCYDCK	240
Db	181	MDSLTFGAPRFLRHLMDSSKKIPVMEFDVAKVLEELTMDQFIDLCILCGDCYDCK	240
QY	241	GGGGTALKLIRHQSGISILENLNKRYQIPEDWPYQEARRLFKEPNVTLDIPELKWTA	300
Db	241	GGGGTALKLIRHQSGISILENLNKRYQIPEDWPYQEARRLFKEPNVTLDIPELKWTA	300
QY	301	PDEGLISFLVKDGNFEDRVTKATEIKSAKNSKSGRLESFPEKPTATTAPLKRKETS	360
Db	301	PDEGLISFLVKDGNFEDRVTKATEIKSAKNSKSGRLESFPEKPTATTAPLKRKETS	360
QY	361	DKTSKAAANKTKAGGKKK 379	

```
|||||
Db 361 DKTSKAAANKTKAGGKKK 379

RESULT 2
US-09-426-557-6
; Sequence 6, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Zea mays
US-09-426-557-6

Query Match 100.0%; Score 379; DB 4; Length 379;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGIGLTKLLADNAPKAMKEQKFESYFGKRIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
Db 1 MGIGLTKLLADNAPKAMKEQKFESYFGKRIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
Qy 61 VTSHLQGMFNRTIRLLEAGIKPVVVEFGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
Db 61 VTSHLQGMFNRTIRLLEAGIKPVVVEFGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
Qy 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLGVPVVEAPSEAECAALCINDKVFVASED 180
Db 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLGVPVVEAPSEAECAALCINDKVFVASED 180
Qy 181 MDSLTFGAPRFLRHLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILCGCDYCDSTK 240
Db 181 MDSLTFGAPRFLRHLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILCGCDYCDSTK 240
Qy 241 GIGGOTALKLIHQHSGSTESILENLNKDRIQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
Db 241 GIGGOTALKLIHQHSGSTESILENLNKDRIQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
Qy 301 PDEGLISFLVKDNGFNEDRV 321
Db 301 PDEGLISFLVKDNGFNEDRV 321

RESULT 3
US-09-426-557-4
; Sequence 4, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Zea mays
US-09-426-557-4

Query Match 73.4%; Score 278; DB 4; Length 379;
Best Local Similarity 99.7%; Pred. No. 2.7e-266;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGIGLTKLLADNAPKAMKEQKFESYFGKRIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
Db 1 MGIGLTKLLADNAPKAMKEQKFESYFGKRIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
Qy 61 VTSHLQGMFNRTIRLLEAGIKPVVVEFGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
Db 61 VTSHLQGMFNRTIRLLEAGIKPVVVEFGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
Qy 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLGVPVVEAPSEAECAALCINDKVFVASED 180
Db 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLGVPVVEAPSEAECAALCINDKVFVASED 180
Qy 181 MDSLTFGAPRFLRHLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILCGCDYCDSTK 240
```

```
; SEQ ID NO 4
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Zea mays
US-09-426-557-4

Query Match 84.7%; Score 321; DB 4; Length 379;
Best Local Similarity 100.0%; Pred. No. 9.8e-309;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGIGLTKLLADNAPKAMKEQKFESYFGKRIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
Db 1 MGIGLTKLLADNAPKAMKEQKFESYFGKRIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
Qy 61 VTSHLQGMFNRTIRLLEAGIKPVVVEFGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
Db 61 VTSHLQGMFNRTIRLLEAGIKPVVVEFGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
Qy 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLGVPVVEAPSEAECAALCINDKVFVASED 180
Db 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLGVPVVEAPSEAECAALCINDKVFVASED 180
Qy 181 MDSLTFGAPRFLRHLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILCGCDYCDSTK 240
Db 181 MDSLTFGAPRFLRHLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILCGCDYCDSTK 240
Qy 241 GIGGOTALKLIHQHSGSTESILENLNKDRIQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
Db 241 GIGGOTALKLIHQHSGSTESILENLNKDRIQIPEDWPYQEARRLFKEPNVTLDIPELKWTA 300
Qy 301 PDEGLISFLVKDNGFNEDRV 321
Db 301 PDEGLISFLVKDNGFNEDRV 321

RESULT 4
US-09-426-557-8
; Sequence 8, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Zea mays
US-09-426-557-8

Query Match 73.4%; Score 278; DB 4; Length 379;
Best Local Similarity 99.7%; Pred. No. 2.7e-266;
Matches 378; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGIGLTKLLADNAPKAMKEQKFESYFGKRIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
Db 1 MGIGLTKLLADNAPKAMKEQKFESYFGKRIAVDASMSIYQFLIVVGRGTGNETLTNEAGE 60
Qy 61 VTSHLQGMFNRTIRLLEAGIKPVVVEFGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
Db 61 VTSHLQGMFNRTIRLLEAGIKPVVVEFGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGD 120
Qy 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLGVPVVEAPSEAECAALCINDKVFVASED 180
Db 121 KDAIEKLSKRTVKVTRQHNEDCKRLRLMLGVPVVEAPSEAECAALCINDKVFVASED 180
Qy 181 MDSLTFGAPRFLRHLMDPSSKKIPVMEFDVAKVLEELTMDQFIDLCILCGCDYCDSTK 240
```

Db 181 KDSLTFGAPRURHLMDSRSSKKIPVMEFDVAKVLEELTMDQFIDICILCGDYCDSTK 240
QY 241 GGGGTALKLIHQHSIESILENKNKRYQIPEDWPYOEARRLFKEPNVTLDIPELKWTA 300
Db 241 GGGGTALKLIHQHSIESILENKNKRYQIPEDWPYOEARRLFKEPNVTLDIPELKWTA 300
QY 301 PDEGLISFLVKGNGFNEDRVTKAIEKTKSAKNKSSQGRLESEFFKPTATTSAFLKRKETS 360
Db 301 PDEGLISFLVKGNGFNEDRVTKAIEKTKSAKNKSSQGRLESEFFKPTATTSAFLKRKETS 360
QY 361 DKTSKAAANKTKAGGKKK 379
Db 361 DKTSKAAANKTKAGGKKK 379

RESULT 5
US-08-455-968E-3
; Sequence 3, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; LENGTH: 377 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-968E-3

Query Match 3.4%; Score 13; DB 2; Length 377;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVYVFDGKPP 91
Db 76 GIKPVYVFDGKPP 88

RESULT 6
US-08-455-968E-10
; Sequence 10, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.

; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-968E-10

Query Match 3.4%; Score 13; DB 2; Length 378;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVYVFDGKPP 91
Db 76 GIKPVYVFDGKPP 88

RESULT 7
US-08-823-516-138
; Sequence 138, Application US/08823516
; Patent No. 5994069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichiev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 138:
SEQUENCE CHARACTERISTICS:
LENGTH: 378 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-823-516-138

Query Match 3.4%: Score 13; DB 2; Length 378;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVYVFDGKPP 91
|||||

DB 76 GIKPVYVFDGKPP 88
|||||

RESULT 8
US-08-455-968E-1
Sequence 1, Application US/08455968E
Patent No. 5874283
GENERAL INFORMATION:
APPLICANT: Harrington, John L.
APPLICANT: Hsieh, Chih-Lin
APPLICANT: Lieber, Michael
TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,968E
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 18985-000100
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-455-968E-1
Query Match 3.4%: Score 13; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 79 GIKPVYVFDGKPP 91
|||||

DB 78 GIKPVYVFDGKPP 90
|||||

RESULT 9
US-08-823-516-137
Sequence 137, Application US/08823516
Patent No. 5994069
GENERAL INFORMATION:
APPLICANT: Hall, Jeff G.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Mast, Andrea L.
APPLICANT: Brow, Mary Ann D.
TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
TITLE OF INVENTION: Sequential Invasive Cleavages
NUMBER OF SEQUENCES: 163
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/823,516
FILING DATE: 24-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/01072
FILING DATE: 21-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/759,038
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/758,314
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/756,386
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02736
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338

```
;
; INFORMATION FOR SEQ ID NO: 137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; US-08-823-516-137

Query Match          3.4%; Score 13; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPYVFDGKPP 91
Db 78 GIKPYVFDGKPP 90

;
; RESULT 10
; US-09-426-557-10
; Sequence 10, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-426-557-10

Query Match          3.4%; Score 13; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. No. 9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPYVFDGKPP 91
Db 78 GIKPYVFDGKPP 90

;
; RESULT 11
; US-08-757-653-176
; Sequence 176, Application US/08757653
; Patent No. 5843669
; GENERAL INFORMATION:
; APPLICANT: Kaiser, Michael W.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Lyamichev, Natasha
; TITLE OF INVENTION: Cleavage Of Nucleic Acid Using
; TITLE OF INVENTION: Thermostable FEN-1 Endonucleases
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757,653
```

```
;
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FOS-02565
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 176:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-757-653-176

Query Match          3.2%; Score 12; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 0.00079;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 EAGIKPVYVFDG 88
Db 70 EAGIKPVYVFDG 81

;
; RESULT 12
; US-08-823-516-79
; Sequence 79, Application US/08823516
; Patent No. 5994069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; TITLE OF INVENTION: Sequential Invasive Cleavages
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01072
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/759,038
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,314
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/756,386
; FILING DATE: 29-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02736
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-823-516-79

Query Match          3.2%; Score 12; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 0.00079;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 EAGIKPVYVFDG 88
Db 70 EAGIKPVYVFDG 81

RESULT 13
US-08-823-516-136
; Sequence 136, Application US/08823516
; Patent No. 5994069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection Of Nucleic Acids By Multiple
; TITLE OF INVENTION: Sequential Invasive Cleavages
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01072
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/759,038
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,314
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/756,386
; FILING DATE: 29-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02736
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 136:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; US-08-823-516-136

Query Match          3.2%; Score 12; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 0.00079;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 EAGIKPVYVFDG 88
Db 70 EAGIKPVYVFDG 81

RESULT 14
US-08-759-038-115
; Sequence 115, Application US/08759038
; Patent No. 6090543
; GENERAL INFORMATION:
; APPLICANT: Prudent, James R.
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Brow, Mary Ann D.
; APPLICANT: Dahlberg, James E.
; TITLE OF INVENTION: Cleavage Of Nucleic Acids
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/759,038
; FILING DATE: 02-DEC-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/
; FILING DATE: 29-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-02574
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 340 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-759-038-115
```

Query Match 3.2%; Score 12; DB 3; Length 340;
Best Local Similarity 100.0%; Pred. No. 0.00079;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 EAGIKPVYVFDG 88
|||||

DB 70 EAGIKPVYVFDG 81
|||||

RESULT 15

US-08-758-314-115
Sequence 115, Application US/08758314
Patent No. 6090606
GENERAL INFORMATION:
APPLICANT: Kaiser, Michael W.
APPLICANT: Lyamichev, Victor I.
APPLICANT: Lyamichev, Natasha
TITLE OF INVENTION: Improved Cleavage Agents
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/758,314
FILING DATE: 02-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/
FILING DATE: 29-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/682,853
FILING DATE: 12-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,491
FILING DATE: 24-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ingolia, Diane E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: FORS-02575
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
LENGTH: 340 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-758-314-115

Query Match 3.2%; Score 12; DB 3; Length 340;
Best Local Similarity 100.0%; Pred. No. 0.00079;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 EAGIKPVYVFDG 88
|||||

DB 70 EAGIKPVYVFDG 81
|||||

RESULT 16

US-08-455-968E-16
Sequence 16, Application US/08455968E
Patent No. 5874283
GENERAL INFORMATION:

APPLICANT: Harrington, John L.
APPLICANT: Hsieh, Chih-Lin
APPLICANT: Lieber, Michael
TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,968E
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 18985-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-455-968E-16

Query Match 2.4%; Score 9; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.045;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 SMSIYQFLI 44
|||||

DB 1 SMSIYQFLI 9
|||||

RESULT 17

US-08-455-968E-5
Sequence 5, Application US/08455968E
Patent No. 5874283
GENERAL INFORMATION:
APPLICANT: Harrington, John L.
APPLICANT: Hsieh, Chih-Lin
APPLICANT: Lieber, Michael
TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,968E
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 382 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-455-968E-5

Query Match 2.4%; Score 9; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 YVFDGKPPD 92
Db 84 YVFDGKPPD 92
|||||

RESULT 18
US-08-823-516-139
; Sequence 139, Application US/08823516
; Patent No. 5994069
; GENERAL INFORMATION:
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Mast, Andrea L.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: Detection of Nucleic Acids By Multiple
; TITLE OF INVENTION: Sequential Invasive Cleavages
; NUMBER OF SEQUENCES: 163
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,516
; FILING DATE: 24-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01072
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/759,038
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/758,314
; FILING DATE: 02-DEC-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/756,386
; FILING DATE: 29-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/682,853
; FILING DATE: 12-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,491
; FILING DATE: 24-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027

; REFERENCE/DOCKET NUMBER: FORS-02736
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 139:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 382 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-823-516-139

Query Match 2.4%; Score 9; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 YVFDGKPPD 92
Db 84 YVFDGKPPD 92
|||||

RESULT 19
US-09-146-319-2
; Sequence 2, Application US/09146319
; Patent No. 6251649
; GENERAL INFORMATION:
; APPLICANT: Matsui, Ikuo
; APPLICANT: Ishikawa, Kazuhiko
; APPLICANT: Kosugi, Yoshitsugu
; APPLICANT: Matsui, Eriko
; APPLICANT: Kawasaki, Satoko
; TITLE OF INVENTION: HYPERSTABLE FLAP ENDONUCLEASE DERIVED FROM A
; TITLE OF INVENTION: HYPERSTABLE FLAP ENDONUCLEASE DERIVED FROM A
; TITLE OF INVENTION: PYROCOCUS
; FILE REFERENCE: 081356-0121
; CURRENT APPLICATION NUMBER: US/09/146,319
; CURRENT FILING DATE: 1998-09-03
; EARLIER APPLICATION NUMBER: JP 239440/1997
; EARLIER FILING DATE: 1997-09-04
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Pyrococcus horikoshii
US-09-146-319-2

Query Match 2.1%; Score 8; DB 4; Length 343;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 YVFDGKPP 91
Db 77 YVFDGKPP 84
|||||

RESULT 20
US-09-175-973-2
; Sequence 2, Application US/09175973A
; Patent No. 6255081
; GENERAL INFORMATION:
; APPLICANT: MATSUI, IKUO
; APPLICANT: ISHIKAWA, KAZUHIKO
; APPLICANT: KOSUGI, YOSHITSUGU
; APPLICANT: MATSUI, ERIKO
; APPLICANT: KAWASAKI, SATOKO
; TITLE OF INVENTION: HYPERSTABLE FLAP ENDONUCLEASE DERIVED FROM A
; TITLE OF INVENTION: HYPERSTABLE FLAP ENDONUCLEASE DERIVED FROM A
; TITLE OF INVENTION: PYROCOCUS
; FILE REFERENCE: 081356/0126
; CURRENT APPLICATION NUMBER: US/09/175,973A
; CURRENT FILING DATE: 1998-10-21

; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Pyrococcus horikoshii
US-09-175-973-2

Query Match 2.1%; Score 8; DB 4; Length 343;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 YVFDGKPP 91
|1111111111|
Db 77 YVFDGKPP 84

RESULT 21

US-08-455-968E-20
; Sequence 20, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Cih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-455-968E-20

Query Match 1.8%; Score 7; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 156 APSEAEA 162
|111111111111|
Db 12 APSEAEA 18

RESULT 22

US-08-858-207A-453
; Sequence 453, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:

; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Glimm, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:

; INFORMATION FOR SEQ ID NO: 453:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6348328e
US-08-858-207A-453

Query Match 1.8%; Score 7; DB 4; Length 117;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 247 ALKLIRQ 253
|111111111111|
Db 35 ALKLIRQ 41

RESULT 23

5463025-4
; Patent No. 5463025
; APPLICANT: Sumi, Yoshihiko; Ichikawa, Yataro; Aoki, Nobuo
; Muramatsu, Masami
; TITLE OF INVENTION: PROTEIN HAVING HUMAN PLASMIN INHIBITING
; ACTIVITY
; NUMBER OF SEQUENCES: 7
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,162
; FILING DATE: 24-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60,691
; FILING DATE: 13-MAY-1993
; APPLICATION NUMBER: 419,913
; FILING DATE: 05-SEP-1989
; SEQ ID NO: 4:
; LENGTH: 179
5463025-4

Query Match 1.8%; Score 7; DB 6; Length 179;

Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 GGQTALK 249
Db 58 GGQTALK 64

RESULT 24
US-09-077-675A-16
; Sequence 16, Application US/09077675A
; Patent No. 6242199
; GENERAL INFORMATION:
; APPLICANT: Pai, Lee-Yuh
; APPLICANT: Feighner, Scott C.
; APPLICANT: Howard, Andrew D.
; APPLICANT: Pong, Sheng-Shung
; APPLICANT: Van Der Ploeg, Leonardus H.T.
; TITLE OF INVENTION: RECEPTOR ASSAY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/077,675A
; FILING DATE: 3-JUN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Cocuzzo, Anna L.
; REGISTRATION NUMBER: 42,452
; REFERENCE/DOCKET NUMBER: 19590P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732-594-1273
; TELEFAX: 732-594-4720
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-077-675A-16

Query Match 1.8%; Score 7; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 EPNVTID 292
Db 11 EPNVTID 17

RESULT 25
5463025-1
; APPLICANT: Sumi, Yoshihiko; Ichikawa, Yataro; Aoki, Nobuo
; Muramatsu, Masami
; TITLE OF INVENTION: PROTEIN HAVING HUMAN PLASMIN INHIBITING
; ACTIVITY
; NUMBER OF SEQUENCES: 7

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,162
; FILING DATE: 24-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60,691
; FILING DATE: 13-MAY-1993
; APPLICATION NUMBER: 419,913
; FILING DATE: 05-SEP-1989
; SEQ ID NO:1:
; LENGTH: 464
5463025-1

Query Match 1.8%; Score 7; DB 6; Length 464;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 GGQTALK 249
Db 31 GGQTALK 37

RESULT 26
US-08-115-365-2
; Sequence 2, Application US/08115365
; Patent No. 5605814
; GENERAL INFORMATION:
; APPLICANT: ABRAMOVITZ, MARK
; APPLICANT: ADAM, MOHAMMED
; APPLICANT: BASTIEN, LISON
; APPLICANT: GRYGORCZYK, RICHARD
; APPLICANT: METTERS, KATHLEEN
; APPLICANT: RUSHMORE, THOMAS H.
; APPLICANT: SAWYER, NICOLE
; TITLE OF INVENTION: DNA ENCODING PROSTAGLANDIN RECEPTOR EP2
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JOHN WALLIN
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/115,365
; FILING DATE: 31-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WALLEN III, JOHN W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-3905
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-115-365-2

Query Match 1.8%; Score 7; DB 1; Length 488;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 KALEKIK 329
Db 339 KALEKIK 345

```
RESULT 27
US-08-586-897-2
; Sequence 2, Application US/08586897
; Patent No. 5759789
; GENERAL INFORMATION:
; APPLICANT: ABRAMOVITZ, MARK
; APPLICANT: ADAM, MOHAMMED
; APPLICANT: BASTIEN, LISON
; APPLICANT: GRYGORCZYK, RICHARD
; APPLICANT: METTERS, KATHLEEN
; APPLICANT: RUSHMORE, THOMAS H.
; APPLICANT: SAWYER, NICOLE
; TITLE OF INVENTION: DNA ENCODING PROSTAGLANDIN RECEPTOR EP2
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JOHN WALLEN
; STREET: P.O. BOX 2000, 126 E. LINCOLN AVE.
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/586,897
; FILING DATE: 05-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/115,365
; FILING DATE: 31-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WALLEN III, JOHN W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19066
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-3905
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-586-897-2
Query Match 1.8%; Score 7; DB 1; Length 488;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 KAEKIK 329
Db 339 KAEKIK 345

RESULT 28
US-08-390-162-6
; Sequence 6, Application US/08390162
; Patent No. 5576192
; GENERAL INFORMATION:
; APPLICANT: Ichikawa, Atsushi
; APPLICANT: Narumiya, Shuh
; TITLE OF INVENTION: Prostaglandin E Receptors, Their DNA and
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
```

```
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,162
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/024179
; FILING DATE: 23-FEB-1993
; APPLICATION NUMBER: JP 036580-1992
; FILING DATE: 24-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 064889-1992
; FILING DATE: 23-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 04221-0020-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 513 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-390-162-6
Query Match 1.8%; Score 7; DB 1; Length 513;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 KAEKIK 329
Db 367 KAEKIK 373

RESULT 29
US-08-685-945B-6
; Sequence 6, Application US/08685945B
; Patent No. 5804415
; GENERAL INFORMATION:
; APPLICANT: Ichikawa, Atsushi
; APPLICANT: Narumiya, Shuh
; TITLE OF INVENTION: Prostaglandin E Receptors, Their DNA and
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,945B
; FILING DATE: 22-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/024179
; FILING DATE: 23-FEB-1993
```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 036580-1992
FILING DATE: 24-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 064889-1992
FILING DATE: 23-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B.
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 04221-0020-02000
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 513 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-685-945B-6

Query Match 1.8%; Score 7; DB 1; Length 513;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 KAIEKIK 329
|||||||
DB 367 KAIEKIK 373

RESULT 30
US-08-588-258B-24
Sequence 24, Application US/08588258B
Patent No. 5929207
GENERAL INFORMATION:
APPLICANT: H. Robert Horvitz et al.
TITLE OF INVENTION: REGULATORS OF G-PROTEIN SIGNALLING
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,258B
FILING DATE: January 12, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bieker-Brady, Kristina
REGISTRATION NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 01997/216001
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 555 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-588-258B-24

Query Match 1.8%; Score 7; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 347 TATTSAP 353
|||||||
DB 292 TATTSAP 298
RESULT 31
US-08-460-505-24
Sequence 24, Application US/08460505
Patent No. 6069296
GENERAL INFORMATION:
APPLICANT: Horvitz, Robert H.
TITLE OF INVENTION: REGULATORS OF G-PROTEIN SIGNALLING
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,505
FILING DATE: 02-JUN-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bieker-Brady, Kristina
REGISTRATION NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 01997/214001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 555 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-505-24

Query Match 1.8%; Score 7; DB 3; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 347 TATTSAP 353
|||||||
DB 292 TATTSAP 298

RESULT 32
PCT-US96-08295-24
Sequence 24, Application PC/TUS9608295
GENERAL INFORMATION:
APPLICANT: Massachusetts Institute of Technology
TITLE OF INVENTION: REGULATORS OF G-PROTEIN SIGNALLING
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: PCT/US96/08295
FILING DATE: 31-MAY-1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/588,258
FILING DATE: 12-JAN-96
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Bieker-Brady, Kristina
REGISTRATION NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 01997/216001
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 555 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-08295-24

Query Match 1.8%; Score 7; DB 5; Length 555;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 347 TATTSAP 353
Db 292 TATTSAP 298

RESULT 33
US-08-557-122A-38
Sequence 38, Application US/08557/122A
Patent No. 5879664
GENERAL INFORMATION:
APPLICANT: Hjort, Carsten Mailand
TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5879664o No. 5879664disk of No. 5879664th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557/122A
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3980.204-US
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 638 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-557-122A-38

Query Match 1.8%; Score 7; DB 2; Length 638;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 AKRYSKR 105
Db 399 AKRYSKR 405

RESULT 34
US-09-262-666-38
Sequence 38, Application US/09262666
Patent No. 6346244
GENERAL INFORMATION:
APPLICANT: Hjort, Carsten Mailand
TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6346244o No. 6346244disk of No. 6346244th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/262.666
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/557,122
FILING DATE: 11-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3980.204-US
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 638 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-262-666-38

Query Match 1.8%; Score 7; DB 4; Length 638;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 AKRYSKR 105
Db 399 AKRYSKR 405

RESULT 35
US-08-557-122A-38
Sequence 38, Application US/08557/122A
Patent No. 5879664
GENERAL INFORMATION:
APPLICANT: Hjort, Carsten Mailand
TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5879664o No. 5879664disk of No. 5879664th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557/122A
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3980.204-US
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 638 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

APPLICATION NUMBER: US/07/587,555
FILING DATE: 24-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 300,211
FILING DATE: 23-JAN-1989
APPLICATION NUMBER: 759,315
FILING DATE: 26-JUL-1985
SEQ ID NO: 5
LENGTH: 664
5240838-5

Query Match 1.8%; Score 7; DB 6; Length 664;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 AGEVTS 64
Db 475 AGEVTS 481

RESULT 36
PCT-US91-09784-4
Sequence 4, Application PC/TUS9109784
GENERAL INFORMATION:
APPLICANT: Davies, Peter JA
APPLICANT: Stein, Joseph P
TITLE OF INVENTION: CLONING AND EXPRESSION OF TISSUE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven W. Parmelee
STREET: One Market Plaza, Steuart Tower, Suite
CITY: San Francisco
STATE: California
COUNTRY: U.S.A.
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/09784
FILING DATE: 19911230
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/635,756
FILING DATE: 04-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 13952-7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 685 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-09784-4

Query Match 1.8%; Score 7; DB 5; Length 685;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 182 DSLTGA 188
Db 55 DSLTGA 61

RESULT 37
US-08-683-262B-75
Sequence 75, Application US/08683262B
Patent No. 5929220
GENERAL INFORMATION:
APPLICANT: Shuping Tong et al.
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/683,262B
FILING DATE: 18-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 31,819
REFERENCE/DOCKET NUMBER: 00786/287002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 973 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-683-262B-75
Query Match 1.8%; Score 7; DB 2; Length 973;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 336 SQRLES 342
Db 122 SQRLES 128

RESULT 38
US-09-361-707-75
Sequence 75, Application US/09361707
Patent No. 6258937
GENERAL INFORMATION:
APPLICANT: Tong, Shuping
Li, Jisu
Wands, Jack R.
TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR
NUMBER OF SEQUENCES: 117
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/361,707

```
;
; FILING DATE: 27-Jul-1999
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/683,262
; FILING DATE: 18-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Creason, Gary L.
; REGISTRATION NUMBER: 34,310
; REFERENCE/DOCKET NUMBER: 00786/287003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 75:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 973 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 75:
US-09-361-707-75
Query Match 1.8%; Score 7; DB 4; Length 973;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 336 SOGRLES 342
DB 122 SOGRLES 128

RESULT 39
US-09-171-461-28
; Sequence 28, Application US/09171461
; Patent No. 6335016
; GENERAL INFORMATION:
; APPLICANT: Baker, Adam
; APPLICANT: Cotten, Matthew
; APPLICANT: Chiocci, Susanna
; APPLICANT: Kurzbauer, Robert
; APPLICANT: Schaffner, Gotthold
; TITLE OF INVENTION: Chicken Embryo Lethal Orphan (CELO) Virus
; FILE REFERENCE: 0652.1800000
; CURRENT APPLICATION NUMBER: US/09/171,461
; CURRENT FILING DATE: 1999-01-12
; EARLIER APPLICATION NUMBER: PCT/EP97/01944
; EARLIER FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 1121
; TYPE: PRT
; ORGANISM: CELO VIRUS
; FEATURE:
; OTHER INFORMATION: Position: 6501..9866/Product: E2b pol
US-09-171-461-28
Query Match 1.8%; Score 7; DB 4; Length 1121;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 KPYYVFD 87
DB 503 KPYYVFD 509

RESULT 40
US-08-061-376-5
; Sequence 5, Application US/08061376
; Patent No. 6175000
; GENERAL INFORMATION:
; APPLICANT: Evans, Glen A.
; APPLICANT: Djabali, Malek
; APPLICANT: Selleri, Lucia
; APPLICANT: Parry, Pauline
; TITLE OF INVENTION: CHARACTERIZATION OF A CHROMOSOME 11Q23
; TITLE OF INVENTION: TRANSLOCATION BREAKPOINT ASSOCIATED WITH ACUTE LEUKEMIAS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/061,376
; FILING DATE: 13-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9387
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)546-4737
; TELEFAX: (619)546-9392
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3969 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-061-376-5
Query Match 1.8%; Score 7; DB 4; Length 3969;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 KRTVKVT 135
DB 2516 KRTVKVT 2522

RESULT 41
PCT-US95-04121-35
; Sequence 35, Application PC/TUS9504121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Haptenated Peptides and Uses Thereof
; NUMBER OF SEQUENCES: 62
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04121
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/222,206
; FILING DATE: April 1, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: VanStone, Darlene A.
; REGISTRATION NUMBER: 35,279
; REFERENCE/DOCKET NUMBER: 079.2PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6010
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
```

;
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
PCT-US95-04121-35

Query Match 1.6%; Score 6; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 364 SKAAN 369
Db 3 SKAAN 8
|||||

RESULT 42
US-08-591-438-12
; Sequence 12, Application US/08591438
; Patent No. 5939063
; GENERAL INFORMATION:
; APPLICANT: Vadas, Mathew A.
; APPLICANT: Lopez, Angel F.
; APPLICANT: Shannon, Mary F.
; TITLE OF INVENTION: HAEMOPOIETIC GROWTH FACTOR ANTAGONISTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/591,438
; FILING DATE: 08-APR-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio Esq., Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9972
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
US-08-591-438-12

Query Match 1.6%; Score 6; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 QEARL 283
Db 6 QEARL 11
|||||

RESULT 43
5405952-5
; Patent No. 5405952
; APPLICANT: DEELEY, MICHAEL; PRICE, VIRGINIA L.; URDAL, DAVID

;
; TITLE OF INVENTION: DNA SEQUENCE ENCODING NONGLYCOSYLATED
; ANALOGS OF HUMAN COLONY STIMULATING FACTORS
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/262,385
; FILING DATE: 24-OCT-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 918,428
; FILING DATE: 14-OCT-1986
; SEQ ID NO: 5:
; LENGTH: 14
5405952-5

Query Match 1.6%; Score 6; DB 6; Length 14;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 QEARL 283
Db 3 QEARL 8
|||||

RESULT 44
US-08-712-212-3
; Sequence 3, Application US/08712212
; Patent No. 5837460
; GENERAL INFORMATION:
; APPLICANT: Williams, William V.
; APPLICANT: Kieber-Emmons, Thomas
; APPLICANT: Weiner, David B.
; APPLICANT: VonFeldt, Joan M.
; TITLE OF INVENTION: Biologically active peptides and
; methods of identifying the same
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESS: No. 5837460ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,212
; FILING DATE: 03-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,404
; FILING DATE: 29-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: OPN-1554
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-712-212-3

Query Match 1.6%; Score 6; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 QEARL 283
|||||
Db 5 QEARL 10

RESULT 45

US-08-712-212-12
; Sequence 12, Application US/08712212
; Patent No. 5837460
; GENERAL INFORMATION:
; APPLICANT: Williams, William V.
; APPLICANT: Kieber-Emons, Thomas
; APPLICANT: Weiner, David B.
; APPLICANT: Vonfeldt, Joan M.
; TITLE OF INVENTION: Biologically active peptides and
; methods of identifying the same
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: NO. 5837460ris
; STREET: One Liberty Place, 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712,212
; FILING DATE: 03-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235,404
; FILING DATE: 29-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: UPN-1554
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-712-212-12

Query Match 1.6% Score 6; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 QEARL 283
|||||
Db 4 QEARL 9

Search completed: November 5, 2002, 16:26:53
Job time : 32 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 16:22:49 ; Search time 45 seconds
(without alignments)
809,286 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 379
Sequence: 1 MG1KGLTKLLADNAPKAMKE.....SDKTSKAAANKTKAGGKKK 379

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR,71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	7.4	362	2 T01198	endonuclease homol
2	15	4.0	343	2 E75117	DNA repair protein
3	14	3.7	385	2 T13692	hypothetical prote
4	13	3.4	378	2 A53730	flap endonuclease
5	13	3.4	380	2 A56531	DNA structure-spec
6	13	3.4	401	2 H72765	probable flap endo
7	12	3.2	340	2 T46893	endo/exonuclease f
8	11	2.9	380	2 A56054	DNA repair endonuc
9	9	2.4	378	2 T41456	probable phosphol
10	9	2.4	382	2 S22267	hypothetical prote
11	8	2.1	107	2 A34006	glutamate--ammonia
12	8	2.1	302	2 F90158	DNA repair endo/ex
13	8	2.1	343	2 A71015	probable 5' nuclea
14	8	2.1	352	2 S10499	recombination prot
15	8	2.1	370	2 S61058	glutamate--ammonia
16	8	2.1	373	1 AJHUQ	glutamate--ammonia
17	8	2.1	373	1 AJHQY	glutamate--ammonia
18	8	2.1	373	1 AJCHO	glutamate--ammonia
19	8	2.1	373	2 S41452	glutamate--ammonia
20	8	2.1	399	1 AJFF1M	glutamate--ammonia
21	8	2.1	559	2 B56731	chromatin assembly
22	7	1.8	63	2 A81064	hypothetical prote
23	7	1.8	67	2 F84040	hypothetical prote
24	7	1.8	96	2 S10069	regulatory protein
25	7	1.8	97	2 B83483	conserved hypothet
26	7	1.8	103	2 B72603	hypothetical prote
27	7	1.8	128	2 C85135	Ribosomal protein
28	7	1.8	133	2 A41298	kinesin-like prote
29	7	1.8	134	2 T50607	hypothetical prote

30	7	1.8	139	2 E81020	conserved hypothet
31	7	1.8	140	2 T13100	probable minor tai
32	7	1.8	145	2 A86490	protein T32E20.11
33	7	1.8	147	1 S24310	flavodoxin - Desul
34	7	1.8	148	2 AB2355	hypothetical prote
35	7	1.8	152	2 H69129	ribosomal protein
36	7	1.8	163	2 B44813	probable molybdenu
37	7	1.8	184	2 D86359	F12K8.6 protein -
38	7	1.8	186	2 D84673	probable calcium b
39	7	1.8	201	2 G72350	conserved hypothet
40	7	1.8	211	2 AB2026	transcription term
41	7	1.8	215	2 D75161	3-dehydroquinat d
42	7	1.8	216	2 C84786	hypothetical prote
43	7	1.8	237	2 A45929	spectrin beta chai
44	7	1.8	242	2 F81363	probable glutamine
45	7	1.8	245	2 F91253	probable DNA modif

ALIGNMENTS

RESULT 1

T01198
endonuclease homolog F21E10.3 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 22-Oct-1999
C:Accession: T01198
R:Davidson, S.; Kohlfing, T.; David, M.; O'Brian, D.
submitted to the EMBL Data Library, April 1998
A:Description: The sequence of A. thaliana F21E10.
A:Reference number: Z14258
A:Accession: T01198
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-362 <D>
A:Cross-references: EMBL:AF058914; NID:g3047074; PID:g3047083; PIDN:AAC13596.1; GSPDB
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ATSP:F21E10.3
A:Map position: 5
A:Introns: 4/3; 51/2; 71/1; 96/2; 116/2; 134/2; 150/3; 166/3; 188/3; 205/3; 245/3; 27
C:Superfamily: conserved hypothetical protein YKL113c
C:Keywords: DNA repair; endonuclease

Query Match 7.4%; Score 28; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 3.7e-20;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 64 HLOGMFNRTIRLEAGIKPVYVFDGKPP 91
Db 97 HLOGMFNRTIRLEAGIKPVYVFDGKPP 124

RESULT 2

E75117
DNA repair protein rad2 (rad2) PAB1877 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: E75117
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s
A:Reference number: A75001
A:Accession: E75117
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-343 <K>
A:Cross-references: GB:AJ248285; GB:AL096836; NID:g5458067; PIDN:CAB49654.1; PID:g545
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1877
C:Superfamily: conserved hypothetical protein YKL113c

```

Query Match          4.0%; Score 15; DB 2; Length 343;
Best Local Similarity 100.0%; Pred. No. 6.6e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 77 EAGIKPVYVFDGKPP 91
   |||||
Db 70 EAGIKPVYVFDGKPP 84

RESULT 3
T13692
hypothetical protein EG0003.3 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C:Accession: T13692
R:Murphy, L.; Harris, D.; Barrell, B.
submitted to the EMBL Data Library, November 1998
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A:Reference number: 217699
A:Accession: T13692
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-385 <MUR>
A:Cross-references: EMBL:AL031863; PIDN:CAA21320.1
C:Genetics:
A:Cross-references: FlyBase:FBgn0025833
A:Introns: 28/1
A>Note: EG:EG0003.3
C:Superfamily: conserved hypothetical protein YKL113C

Query Match          3.7%; Score 14; DB 2; Length 385;
Best Local Similarity 100.0%; Pred. No. 7.7e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 GIKPVYVFDGKPPD 92
   |||||
Db 78 GIKPVYVFDGKPPD 91

RESULT 4
A53730
flap endonuclease 1 - mouse
N:Alternate names: FEN-1
C:Species: Mus musculus (house mouse)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999
C:Accession: A53730
R:Harrington, J.J.; Lieber, M.R.
Genes Dev. 8, 1344-1355, 1994
A:Title: Functional domains within FEN-1 and RAD2 define a family of structure-specific
A:Reference number: A53730; MUID:95011546
A:Accession: A53730
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-378 <HAR>
A:Cross-references: GB:L26320; NID:g499648; PIDN:AAC37664.1; PID:g499649
C:Superfamily: conserved hypothetical protein YKL113C
C:Keywords: DNA repair; endonuclease

Query Match          3.4%; Score 13; DB 2; Length 378;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 GIKPVYVFDGKPP 91
   |||||
Db 76 GIKPVYVFDGKPP 88

RESULT 5
A56531
DNA structure-specific endonuclease FEN1 [validated] - human
N:Alternate names: flap endonuclease 1; rad2 protein homolog
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 21-Jul-2000

```

```

C:Accession: A56531; B56054
R:Hiraoka, L.R.; Harrington, J.J.; Gerhard, D.S.; Lieber, M.R.; Hsieh, C.L.
Genomics 25, 220-225, 1995
A:Title: Sequence of human FEN-1, a structure-specific endonuclease and chromosomal
A:Reference number: A56531; MUID:95293376
C:Accession: A56531
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-380 <HIR>
A:Cross-references: GB:L37374; NID:g642089; PIDN:AAA91331.1; PID:g642090
R:Murray, J.M.; Tavassoli, M.; Al-Harithy, R.; Sheldrick, K.S.; Lehmann, A.R.; Carr,
Mol. Cell. Biol. 14, 4878-4888, 1994
A:Title: Structural and functional conservation of the human homolog of the Schizosac
A:Reference number: A56054; MUID:94277093
C:Accession: B56054
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-380 <MUR>
A:Cross-references: GB:S70527
C:Genetics:
A:Gene: GDB:FEN1
A:Cross-references: GDB:451178; OMIM:600393
A:Map position: 11q12-11q12
C:Function:
A:Description: cleaves substrates with unannealed 5'-tails [validated, MUID:20209428]
A>Note: needed during Okazaki fragment processing and long patch base excision repair
C:Superfamily: conserved hypothetical protein YKL113C
C:Keywords: DNA repair; endonuclease

Query Match          3.4%; Score 13; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 GIKPVYVFDGKPP 91
   |||||
Db 78 GIKPVYVFDGKPP 90

RESULT 6
H72765
probable flap endonuclease-1 APE0115 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: H72765
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339
A:Accession: H72765
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-401 <KAW>
A:Cross-references: DDBJ:AP000058; NID:g5103388; PIDN:BAA79026.1; PID:g5103505
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0115
C:Superfamily: conserved hypothetical protein YKL113C

Query Match          3.4%; Score 13; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 8.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 79 GIKPVYVFDGKPP 91
   |||||
Db 126 GIKPVYVFDGKPP 138

RESULT 7
T46893
endo/exonuclease fen-1 [imported] - Pyrococcus furiosus
N:Alternate names: DNA repair protein
C:Species: Pyrococcus furiosus

```

C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 21-Jul-2000
 C:Accession: T46893
 R:DiRuggiero, J.; Brown, J.R.; Bogert, A.P.; Robb, F.T.
 J. Mol. Evol. 49, 474-484, 1999
 A:Title: DNA repair systems in Archaea: mementos from the last universal common ancestor
 A:Reference number: 224126; MUID:99415851
 A:Accession: T46893
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-340 <DR>
 A:Cross-references: EMBL:AF013497; NID:g4102601; PIDN:AAD01514.1; PID:g4102602
 A:Experimental source: strain DSM 3638
 C:Genetics:
 A:Gene: fen-1
 C:Superfamily: conserved hypothetical protein YKL113c

Query Match 3.2%; Score 12; DB 2; Length 340;
 Best Local Similarity 100.0%; Pred. No. 0.00076;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 EAGIKPVYVFDG 88
 |||||
 Db 70 EAGIKPVYVFDG 81

RESULT 8
 A56054
 DNA repair endonuclease rad2 - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 21-Jul-2000
 A:Accession: A56054; T45210; T38725
 R:Murray, J.M.; Tavassoli, M.; Al-Harithy, R.; Sheldrick, K.S.; Lehmann, A.R.; Carr, A.M.
 Mol. Cell. Biol. 14, 4878-4888, 1994
 A:Title: Structural and functional conservation of the human homolog of the Schizosaccharomyces pombe protein YKL113c
 A:Reference number: A56054; MUID:94277093
 A:Accession: A56054
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-380 <MUR>
 A:Cross-references: GB:S70605
 A:Accession: T45210
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Residues: 1-380 <MU2>
 A:Cross-references: EMBL:X77041; NID:g2909339; PIDN:CAB36991.1; PID:g4456854
 A:Experimental source: strain 972h(-)
 R:Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, September 1997
 A:Reference number: 221797
 A:Accession: T38725
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-380 <GEN>
 A:Cross-references: EMBL:Z99167; PIDN:CAB16282.1; GSPDB:GN000666; SPDR:SPAC3G6.06c
 A:Experimental source: strain 972h; cosmid c3G6
 C:Genetics:
 A:Gene: rad2
 A:Map position: 1
 A:Introns: 5/1; 35/2
 C:Function:
 A:Description: DNA repair
 C:Superfamily: conserved hypothetical protein YKL113c

Query Match 2.9%; Score 11; DB 2; Length 380;
 Best Local Similarity 100.0%; Pred. No. 0.0089;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 KTKVKVTRQHN 139
 |||||
 Db 129 KTKVKVTRQHN 139

RESULT 9

T41456
 probable phospholipase SPCC5E4.05c - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 07-Dec-1999
 A:Accession: T41456
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
 submitted to the EMBL Data Library, November 1998
 A:Reference number: 221995
 A:Accession: T41456
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-378 <WOO>
 A:Cross-references: EMBL:AL033406; PIDN:CAA21960.1; GSPDB:GN00068; SPDB:SPCC5E4.05c
 A:Experimental source: strain 972h; cosmid c5E4
 C:Genetics:
 A:Gene: SPDB:SPCC5E4.05c
 A:Map position: 3

Query Match 2.4%; Score 9; DB 2; Length 378;
 Best Local Similarity 100.0%; Pred. No. 0.97;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 345 KPTATTAP 353
 |||||
 Db 292 KPTATTAP 300

RESULT 10
 S22267
 hypothetical protein YKL113c - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein YKL510; RAD2 protein homolog
 C:Species: Saccharomyces cerevisiae
 C:Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 24-Sep-1999
 C:Accession: S22267; S37941
 R:Jacquier, A.; Legrain, P.; Dujon, B.
 Yeast 8, 121-132, 1992
 A:Title: Sequence of a 10.7 kb segment of yeast chromosome XI identifies the APN1 and S.
 A:Reference number: S22267; MUID:92221689
 A:Accession: S22267
 A:Molecule type: DNA
 A:Residues: 1-382 <JAC>
 A:Cross-references: GB:S93804; NID:g248391; PIDN:AAB21998.1; PID:g248392
 A:Experimental source: strain S288C
 R:Jacquier, A.; Legrain, P.; Colletaux, L.; Richard, G.F.; Thierry, A.; Dujon, B.
 submitted to the Protein Sequence Database, March 1994
 A:Reference number: S37938
 A:Accession: S37941
 A:Molecule type: DNA
 A:Residues: 1-382 <JAC>
 A:Cross-references: EMBL:228113; NID:g486189; PIDN:CAA81953.1; PID:g486190; MIPS:YKL1
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGB:RAD27
 A:Cross-references: SGD:S0001596; MIPS:YKL113c
 A:Map position: 11L
 C:Superfamily: conserved hypothetical protein YKL113c
 C:Keywords: nucleus

Query Match 2.4%; Score 9; DB 2; Length 382;
 Best Local Similarity 100.0%; Pred. No. 0.96;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 YVFDGKPPD 92
 |||||
 Db 84 YVFDGKPPD 92

RESULT 11
 A34006
 glutamate--ammonia ligase (EC 6.3.1.2) - bovine (fragment)
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 16-Jul-1999

C:Accession: A34006
 R:Smith Jr., D.D.; Campbell, J.W.
 Proc. Natl. Acad. Sci. U.S.A. 85, 160-164, 1988
 A:Title: Distribution of glutamine synthetase and carbamoyl-phosphate synthetase I in *Escherichia coli*
 A:Reference number: A34006; MUID:88124783
 A:Accession: A34006
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-107 <SMI>
 A:Cross-references: GB:J03604; NID:g163100; PIDN:AAA87357.1; PID:g163101
 C:Superfamily: glutamate--ammonia ligase
 C:Keywords: ligase

Query Match 2.1%; Score 8; DB 2; Length 107;
 Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 ATEKLSKR 130
 |||||
 Db 7 ATEKLSKR 14

RESULT 12
 F90158
 DNA repair endo/exonuclease FEN-1 (RAD2) (rad2) [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
 C:Accession: F90158
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Avayez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
 arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: F90158
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-302 <KUP>
 A:Cross-references: GB:AE006641; NID:gl3813313; PIDN:AAK40525.1; GSPDB:GN00155
 C:Gene: rad2
 C:Genetics:
 C:Superfamily: conserved hypothetical protein YKL113c

Query Match 2.1%; Score 8; DB 2; Length 302;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 YVFDGKPP 91
 |||||
 Db 30 YVFDGKPP 37

RESULT 13
 A71015
 probable 5' nuclease - Pyrococcus horikoshii
 C:Species: Pyrococcus horikoshii
 C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
 C:Accession: A71015
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki-
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi,
 DNA Res. 5, 55-76, 1998
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
 A:Reference number: A71000; MUID:98344137
 A:Accession: A71015
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-343 <KAW>
 A:Cross-references: GB:AP000006; MID:g3236133; PIDN:BAAS0521.1; PID:g3257838
 A:Experimental source: strain OT3
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C:Genetics:
 A:Gene: PH1415
 C:Superfamily: conserved hypothetical protein YKL113c

Query Match 2.1%; Score 8; DB 2; Length 343;
 Best Local Similarity 100.0%; Pred. No. 9.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 YVFDGKPP 91
 |||||
 Db 77 YVFDGKPP 84

RESULT 14
 SI0499
 recombination protein recA - Bordetella pertussis
 N:Alternate names: recombinase A
 C:Species: Bordetella pertussis
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001
 C:Accession: SI0499
 R:Favre, D.; Viret, J.F.
 Nucleic Acids Res. 18, 4243, 1990
 A:Title: Nucleotide sequence of the recA gene of Bordetella pertussis.
 A:Reference number: SI0499; MUID:90332418
 A:Accession: SI0499
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-352 <FAV>
 A:Cross-references: EMBL:X53457
 C:Genetics:
 A:Gene: recA
 C:Superfamily: recombination protein recA
 C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-lo
 F73-80/Region: nucleotide-binding motif A (P-loop)
 F:147-152/Region: nucleotide-binding motif B
 F:79/Binding site: ATP (lys) #status predicted

Query Match 2.1%; Score 8; DB 2; Length 352;
 Best Local Similarity 100.0%; Pred. No. 9.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 361 DKTSKAAA 368
 |||||
 Db 3 DKTSKAAA 10

RESULT 15
 S61058
 glutamate--ammonia ligase (EC 6.3.1.2) - yeast (Saccharomyces cerevisiae)
 N:Alternate names: glutamine synthetase; protein YP3085.01; protein YP9367.15; protei
 C:Species: Saccharomyces cerevisiae
 C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 16-Jul-1999
 C:Accession: S61058; A42278; S36939; A28519; S54509
 R:Badcock, K.; Churcher, C.M.
 submitted to the EMBL Data Library, November 1995
 A:Reference number: S61058
 A:Accession: S61058
 A:Molecule type: DNA
 A:Residues: 1-370 <BAD>
 A:Cross-references: EMBL:Z68111; NID:g1072402; PIDN:CAA92141.1; PID:g1072403; MIPS:YP
 R:Minohart, P.L.; Magasanik, B.
 J. Bacteriol. 174, 1828-1836, 1992
 A:Title: Sequence of the GUN1 gene of Saccharomyces cerevisiae: role of the upstream
 A:Reference number: A42278; MUID:92193267
 A:Accession: A42278
 A:Molecule type: DNA
 A:Residues: 1-250, 'T', 252-263, 'M', 265-370 <MIN>
 A:Cross-references: EMBL:M65157
 A:Note: sequence extracted from NCBI backbone (NCBIN:88660, NCBIP:88667)
 R:Magasanik, B.; Minohart, P.
 submitted to the EMBL Data Library, May 1992
 A:Reference number: S36939
 A:Accession: S36939
 A:Molecule type: DNA
 A:Residues: 1-44, 69-250, 'T', 252-263, 'M', 265-370 <MAG>
 A:Cross-references: EMBL:M65157; NID:g171597; PIDN:AAA34644.1; PID:g171598

R:Kim, K.H.; Rhee, S.G.
 J. Biol. Chem. 263, 833-838, 1988
 A:Title: Sequence of peptides from Saccharomyces cerevisiae glutamine synthetase. N-term
 A:Reference number: A28519; MUID:88087208
 A:Accession: A28519
 A:Molecule type: protein
 A:Residues: 2-8;149-164,'D',166;171,'V',173-178;220-224;266-273 <KIM>
 R:Badcock, K.; Churcher, C.M.
 submitted to the EMBL Data Library, May 1995
 A:Reference number: S54509
 A:Accession: S54509
 A:Molecule type: DNA
 A:Residues: 1-171 <BAW>
 A:Cross-references: EMBL:Z49274; NID:g809585; PIDN:CAA89289.1; PID:g809600; MIPS:YPR035W
 A:Experimental source: strain AB972
 C:Genetics:
 A:Gene: SGD:GLM1
 A:Cross-references: SGD:S00062239; MIPS:YPR035W
 A:Map position: 16r
 C:Superfamily: glutamate--ammonia ligase
 C:Keywords: ligase

Query Match 2.1%; Score 8; DB 2; Length 370;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
 |||||
 DB 270 AIEKLSKR 277

RESULT 16
 AJHUQ
 glutamate--ammonia ligase (EC 6.3.1.2) - human
 N:Alternate names: glutamine synthetase
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1991 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
 C:Accession: S18455; S00673; I53634; T47179; S15485
 R:van den Hoff, M.J.B.; Geerts, W.J.C.; Das, A.T.; Moorman, A.F.M.; Lamers, W.H.
 Biochim. Biophys. Acta 1090, 249-251, 1991
 A:Title: cDNA sequence of the long mRNA for human glutamine synthetase.
 A:Reference number: S18455; MUID:92031701
 A:Accession: S18455
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-373 <HOF>
 A:Cross-references: EMBL:X59834; NID:g31830; PIDN:CAA42495.1; PID:g31831
 R:Gibbs, C.S.; Campbell, K.E.; Wilson, R.H.
 Nucleic Acids Res. 15, 6293, 1987
 A:Title: Sequence of a human glutamine synthetase cDNA.
 A:Reference number: S00673; MUID:87316885
 A:Accession: S00673
 A:Molecule type: mRNA
 A:Residues: 1-313,'G',315-321,'SI',324-346,'D',348-373 <GIB>
 A:Cross-references: EMBL:Y00387; NID:g31832; PIDN:CAA68457.1; PID:g31833
 R:Christa, L.; Simon, M.T.; Flinois, J.P.; Gebhardt, R.; Brechot, C.; Lasserre, C.
 Gastroenterology 106, 1312-1320, 1994
 A:Title: Overexpression of glutamine synthetase in human primary liver cancer.
 A:Reference number: I53634; MUID:94229472
 A:Accession: I53634
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-321,'SI',324-346,'D',348-373 <RES>
 A:Cross-references: GB:S70290; NID:g546602; PIDN:AAB30693.1; PID:g546603
 R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Weil, B.; Wiemann, S.
 submitted to the Protein Sequence Database, March 2000
 A:Reference number: Z24378
 A:Accession: T47179
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-109,'GVL' <AAA>
 A:Cross-references: EMBL:AL161952
 A:Experimental source: adult testis; clone DKFp434M0813

C:Genetics:
 A:Gene: GDB:GLUL; GLNS
 A:Cross-references: GDB:I20626; OMIM:138290
 A:Map position: lq31-lq31
 A:Note: DKFp434M0813.1
 C:Superfamily: glutamate--ammonia ligase
 C:Keywords: ligase

Query Match 2.1%; Score 8; DB 1; Length 373;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
 |||||
 DB 273 AIEKLSKR 280

RESULT 17
 AJHUQ
 glutamate--ammonia ligase (EC 6.3.1.2) - Chinese hamster
 N:Alternate names: glutamine synthetase
 C:Species: Cricetulus griseus (Chinese hamster)
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 30-Jun-1993
 C:Accession: A23524
 R:Hayward, B.E.; Hussain, A.; Wilson, R.H.; Lyons, A.; Woodcock, V.; McIntosh, B.; Ha
 Nucleic Acids Res. 14, 999-1008, 1986
 A:Title: The cloning and nucleotide sequence of cDNA for an amplified glutamine synth
 A:Reference number: A23524; MUID:86120385
 A:Accession: A23524
 A:Molecule type: mRNA
 A:Residues: 1-373 <HAY>
 A:Cross-references: GB:X03459
 C:Superfamily: glutamate--ammonia ligase
 C:Keywords: ligase

Query Match 2.1%; Score 8; DB 1; Length 373;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
 |||||
 DB 273 AIEKLSKR 280

RESULT 18
 AJCHQ
 glutamate--ammonia ligase (EC 6.3.1.2) - chicken
 N:Alternate names: glutamine synthetase
 C:Species: Gallus gallus (chicken)
 C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1999
 C:Accession: JQ0025; A44095
 R:Fu, H.; Young, A.P.
 Gene 81, 169-175, 1989
 A:Title: The structure of the chicken glutamine synthetase-encoding gene.
 A:Reference number: JQ0025; MUID:90034178
 A:Accession: JQ0025
 A:Molecule type: mRNA
 A:Residues: 1-373 <PUH>
 A:Cross-references: GB:M29076; NID:g211814; PIDN:AAA48783.1; PID:g211815
 R:Campbell, J.W.; Smith Jr., D.D.
 Mol. Biol. Evol. 9, 787-805, 1992
 A:Title: Metabolic compartmentation of vertebrate glutamine synthetase: putative mito
 A:Reference number: A44095; MUID:92408449
 A:Accession: A44095
 A:Molecule type: mRNA
 A:Residues: 1-373 <CAM>
 A:Cross-references: GB:S45408; NID:g255851; PIDN:AAC69361.1; PID:g255852
 A:Experimental source: liver
 A:Note: sequence extracted from NCBI backbone (NCBIP:113559)
 C:Superfamily: glutamate--ammonia ligase
 C:Keywords: ligase

Query Match 2.1%; Score 8; DB 1; Length 373;

Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
Db 273 AIEKLSKR 280

RESULT 19
S41452
glutamate--ammonia ligase (EC 6.3.1.2) - pig
A:Alternate names: glutamine synthetase
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Jul-1999
C:Accession: S41452
R:Johnstone, R.W.; Loveland, B.E.
submitted to: The EMBL Data Library, January 1994
A:Description: The cloning and nucleotide sequence of porcine glutamine synthetase cDNA.
A:Reference number: S41452
A:Accession: S41452
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-373 <JOH>
A:Cross-references: EMBL:Z29636; NID:g452436; PIDN:CAA82747.1; PID:g452437.
C:Superfamily: glutamate--ammonia ligase
C:Keywords: ligase

Query Match 2.1%; Score 8; DB 2; Length 373;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
Db 273 AIEKLSKR 280

RESULT 20
AJFFIM
glutamate--ammonia ligase (EC 6.3.1.2) 1, mitochondrial - fruit fly (Drosophila melanogaster)
A:Alternate names: glutamine synthetase
C:Species: Drosophila melanogaster
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Feb-1997
C:Accession: S09109
R:Caizzi, R.; Bozzetti, M.P.; Caggese, C.
J. Mol. Biol. 212, 17-26, 1990
A:Title: Homologous nuclear genes encode cytoplasmic and mitochondrial glutamine synthetase
A:Reference number: S09108; MUID:90204539
A:Accession: S09109
A:Molecule type: DNA
A:Residues: 1-399 <CAI>
C:Genetics:
A:Gene: FlyBase:Gsl
A:Cross-references: FlyBase:FBgn0001142
C:Superfamily: glutamate--ammonia ligase
C:Keywords: ligase; mitochondrion

Query Match 2.1%; Score 8; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
Db 309 AIEKLSKR 316

RESULT 21
B56731
chromatin assembly factor I p60 chain - human
C:Species: Homo sapiens (man)
C:Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 21-Jul-2000
C:Accession: B56731
R:Kaufman, P.D.; Kobayashi, R.; Kessler, N.; Stillman, B.
Cell 81, 1105-1114, 1995

A:Title: The p150 and p60 subunits of chromatin assembly factor I: a molecular link b
A:Reference number: A56731; MUID:95323966
A:Accession: B56731
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-559 <RAU>
A:Cross-references: GB:U20980; NID:g882259; PIDN:AAA76737.1; PID:g882260
C:Genetics:
A:Gene: GDB:CAFLA
A:Cross-references: GDB:728461; OMIM:601245
A:Map position: 21q22.2-21q22.3
C:Superfamily: unassigned WD repeat proteins; WD repeat homology
F:62-95/Domain: WD repeat homology <WD2>
F:125-158/Domain: WD repeat homology <WD3>

Query Match 2.1%; Score 8; DB 2; Length 559;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 VFAVASED 180
Db 318 VFAVASED 325

RESULT 22
A81064
hypothetical protein NMB1598 [imported] - Neisseria meningitidis (strain MC58 serogro
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: A81064
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, B.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: A81064
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-63 <TET>
A:Cross-references: GB:AE002511; GB:AE002098; NID:g7226857; PIDN:AAF41951.1; PID:g722
C:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1598

Query Match 1.8%; Score 7; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 KFESYFG 28
Db 4 KFESYFG 10

RESULT 23
F84040
hypothetical protein BH3126 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: F84040
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: F84040
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-67 <STO>
A:Cross-references: GB:AP001517; GB:BA000004; NID:g10175500; PIDN:BAF06845.1; GSPDB:G
C:Experimental source: strain C-125
C:Genetics:
A:Gene: BH3126


```

Query Match          1.8%; Score 7; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 KAIEKIK 329
      |||||
DB 14 KAIEKIK 20

RESULT 24
SI0069
regulatory protein N - phage P22
C:Species: phage P22
C:Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 08-Oct-1999
C:Accession: SI0069
J. Franklin, N.C.
R. Mol. Biol. 181, 75-84, 1985
A:Title: Conservation of genome form but not sequence in the transcription antitermination
A:Reference number: S07268; MUID:85160834
A:Accession: SI0069
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-96 <FRA>
A:CROSS-references: EMBL:X02141; NID:g15644; PIDN:CRA26079.1; PID:g15645
C:Genetics:
A:Gene: N
C:Keywords: transcription regulation

Query Match          1.8%; Score 7; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 DRVTKAI 325
      |||||
DB 58 DRVTKAI 64

RESULT 25
B83483
conserved hypothetical protein PA1295 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83483
R.Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizooguchi, S.D.; Warrenner, P.; Hickey, M.J.; Ba
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A:Reference number: A82950; MUID:20437337
A:Accession: B83483
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-97 <STO>
A:CROSS-references: GB:AE004559; GB:AE004091; NID:g9947228; PIDN:AAG04684.1; GSPDB:GN001
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA1295

Query Match          1.8%; Score 7; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 DVAKVLE 215
      |||||
DB 58 DVAKVLE 64

RESULT 26
B72603
hypothetical protein APE1289 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000

C:Accession: B72603
R.Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339
A:Accession: B72603
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-103 <KAW>
A:CROSS-references: DDBJ:AP000061; NID:g5104821; PIDN:BAA80280.1; PID:d1044066; PID:g
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1289
C:Superfamily: Aeropyrum pernix hypothetical protein APE1289

Query Match          1.8%; Score 7; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 GVPVVEA 156
      |||||
DB 91 GVPVVEA 97

RESULT 27
C85135
Ribosomal protein L7Ae-like [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: C85135
R.anonymous. The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488
A:Accession: C85135
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-128 <STO>
A:CROSS-references: GB:NC_001268; NID:g7267962; PIDN:CAB78303.1; GSPDB:GN00140
A:Genetics:
A:Gene: At4g12600
A:Map position: 4

Query Match          1.8%; Score 7; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 KDAIEKL 127
      |||||
DB 120 KDAIEKL 126

RESULT 28
A41298
kinesin-like protein 1 - fruit fly (Drosophila melanogaster) (fragment)
C:Species: Drosophila melanogaster
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 24-Sep-1999
C:Accession: A41298
R.Stewart, R.J.; Pesavento, P.A.; Woerpel, D.N.; Goldstein, L.S.B.
Proc. Natl. Acad. Sci. U.S.A. 88, 8470-8474, 1991
A:Title: Identification and partial characterization of six members of the kinesin su
A:Reference number: A41298; MUID:92020874
A:Accession: A41298
A:Molecule type: DNA
A:Residues: 1-133 <STE>
A:CROSS-references: GB:M74427; NID:g157783; PIDN:AAA28654.1; PID:g157784
C:Genetics:
A:Gene: FlyBase:Klp54D
A:CROSS-references: FlyBase:FBgn0004377
C:Superfamily: unassigned kinesin-related proteins: kinesin motor domain homology
C:Keywords: ATP; P-loop
F:1-133/domain: kinesin motor domain homology (fragment) <KNOT>

```

Query Match 1.8%; Score 7; DB 2; Length 133;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LTKLLAD 12
 |||||
 Db 89 LTKLLAD 95

RESULT 29

T50607
 hypothetical protein DKFzp43411016.1 - human (fragment)

C:Species: Homo sapiens (man)
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000

C:Accession: T50607
 R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Weil, B.; Wiemann, S.

submitted to the Protein Sequence Database, June 2000

A:Reference number: 225141

A:Accession: T50607

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-134 <AAA>

A:Cross-references: EMBL:AL359560

A:Experimental source: adult testis; clone DKFzp43411016

C:Genetics:

A:Note: DKFzp43411016.1

Query Match 1.8%; Score 7; DB 2; Length 134;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 352 APLKRKE 358
 |||||
 Db 22 APLKRKE 28

RESULT 30

E81020

conserved hypothetical protein NMB1980 [imported] - Neisseria meningitidis (strain MC58)

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C:Accession: E81020

R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haif, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

Li, H.; Qin, H.; Vamathavan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755

A:Accession: E81020

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-139 <TET>

A:Cross-references: GB:AE002546; GB:AE002098; NID:g7227234; PIDN:AAF42308.1; PID:g722724

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB1980

Query Match 1.8%; Score 7; DB 2; Length 139;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 AALCIND 171
 |||||
 Db 85 AALCIND 91

RESULT 31

T13100

probable minor tail protein G - phage N15

N:Alternate names: tail assembly protein gp14

C:Species: phage N15

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 13-Aug-1999
 C:Accession: T13100
 R:Hendrix, R.W.; Ravin, V.K.; Casjens, S.R.; Ford, M.E.; Ravin, N.V.; Smirnov, I.K.

submitted to the EMBL Data Library, May 1998

A:Reference number: Z17603

A:Accession: T13100

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-140 <HEN>

A:Cross-references: EMBL:AF064539; NID:g3192683; PID:g3192692; PIDN:AAC19045.1

C:Genetics:

A:Note: gene 14

C:Superfamily: phage lambda minor tail protein G

C:Keywords: tail protein

Query Match 1.8%; Score 7; DB 2; Length 140;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 114 EAVEVGD 120
 |||||
 Db 36 EAVEVGD 42

RESULT 32

A86490

protein T32E20.11 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: A86490

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, Z.A.; Luros, J.S.; Maifi, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: A86490

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-146 <SPO>

A:Cross-references: GB:AE005172; NID:g8778775; PIDN:AAF79783.1; GSPDB:GN00141

C:Genetics:

A:Gene: T32E20.11

A:Map position: 1

Query Match 1.8%; Score 7; DB 2; Length 146;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 EFDVAKV 213
 |||||
 Db 119 EFDVAKV 125

RESULT 33

S24310

flavodoxin - Desulfovibrio gigas (ATCC 29494)

C:Species: Desulfovibrio gigas

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 05-May-2000

C:Accession: S24310

R:Helms, L.R.; Swenson, R.P.

Biochim. Biophys. Acta 1131, 325-328, 1992

A:Title: The primary structures of the flavodoxins from two strains of Desulfovibrio

A:Reference number: S24310; MUID:92329549

A:Accession: S24310

A:Molecule type: DNA

A:Residues: 1-147 <HEL>

A:Cross-references: EMBL:X64765; NID:g40798; PIDN:CAA46012.1; PID:g40799

A:Experimental source: strain ATCC 29494
 C:Superfamily: flavodoxin; flavodoxin homology
 C:Keywords: electron transfer; flavoprotein; FMN
 F:6-143/Domain: flavodoxin homology <FLX>

Query Match 1.8%; Score 7; DB 1; Length 147;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 VAKVLEE 216
 |||||

DB 22 VAKVLEE 28

RESULT 34
 AB2355

hypothetical protein alr4394 [imported] - Anabaena sp. (strain PCC 7120)

A:Species: Anabaena sp.
 A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
 C:Accession: AB2355

R.Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriquchi, S.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AB2355
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-148 <KUR>

A:CROSS-references: GB:BA000019; PIDN:BA076093.1; PID:gl7133530; GSPDB:GN00179
 A:Experimental source: strain PCC 7120

C:Genetics:
 A:Gene: alr4394

Query Match 1.8%; Score 7; DB 2; Length 148;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 KGLTKLL 10
 |||||

DB 109 KGLTKLL 115

RESULT 35
 H69129

ribosomal protein L30 - Methanobacterium thermoautotrophicum (strain Delta H)
 N:Alternate names: eukaryotic ribosomal protein L7; prokaryotic ribosomal protein L30 hc
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 13-Aug-1999

C:Accession: H69129

R.Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Olu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.; Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: function
 A:Reference number: A69000; MUID:98037514

A:Accession: H69129

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-152 <MTH>

A:CROSS-references: GB:AF000796; GB:AE000666; NID:g2621057; PIDN:AAB84533.1; PID:g262106
 A:Experimental source: strain Delta H

C:Genetics:
 A:Gene: MTH24

C:Superfamily: rat ribosomal protein L7

Query Match 1.8%; Score 7; DB 2; Length 152;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 EAGIKPV 83

|||||

Db 108 EAGIKPV 114

RESULT 36
 B44813

probable molybdenum cofactor biosynthesis protein nifB [similarity] - Herbaspirillum
 C:Species: Herbaspirillum seropedicae
 C:Date: 31-Mar-1993 #sequence_revision 26-Apr-1996 #text_change 02-Jun-2000
 C:Accession: B44813
 R.Souza, E.M.; Funayama, S.; Rigo, L.U.; Yates, M.G.; Pedrosa, F.O.
 J. Gen. Microbiol. 137, 1511-1522, 1991

A:Title: Sequence and structural organization of a nifA-like gene and part of a nifB-
 A:Reference number: A44813; MUID:92065220

A:Accession: B44813
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-163 <SOU>

A:Experimental source: strain 278
 A:Note: sequence extracted from NCBI backbone (NCBIN:56814, NCBIP:56816)

C:Superfamily: Rhizobium nifB protein

Query Match 1.8%; Score 7; DB 2; Length 163;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 KVFAVAS 178

|||||

DB 115 KVFAVAS 121

RESULT 37
 D86359

F12K8.6 protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Nov-2001

C:Accession: D86359

R.Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: D86359

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-184 <STO>

A:CROSS-references: GB:AE005172; NID:g6587850; PIDN:AAF18539.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

C:Superfamily: Arabidopsis hypothetical protein F19F18.60

Query Match 1.8%; Score 7; DB 2; Length 184;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 VYVFDGK 89

|||||

DB 50 VYVFDGK 56

RESULT 38
 D84673

probable calcium binding protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: D84673

R.Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A:Reference-number: A84420; MUID:20083487
A:Accession: D84673
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-186 <STO>
A:Cross-references: GB:AE002093; NID:g4314390; PIDN:AADI5600.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g27480
A:Map position: 2

Query Match 1.8%; Score 7; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 NRTIRLL 76
|||||
Db 44 NRTIRLL 50

RESULT 39
G72350
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: G72350
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing of *Thermotoga maritima*.
A:Reference number: A72200; MUID:99287316
A:Accession: G72350
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-201 <ARN>
A:Cross-references: GB:AE001738; GB:AE000512; NID:g4981158; PIDN:AAD35724.1; PID:g498116
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0640
C:Superfamily: Alkaligenes eutrophus phosphoglycolate phosphatase

Query Match 1.8%; Score 7; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 209 DVAKVLE 215
|||||
Db 29 DVAKVLE 35

RESULT 40
AB2026
transcription termination factor nusB [imported] - Anabaena sp. (strain PCC 7120)
C:Species: Anabaena sp.
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AB2026
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anabaena* sp. strain PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB2026
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-211 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA073459.1; PID:g17130850; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: nusB

C:Keywords: transcription termination

Query Match 1.8%; Score 7; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 ELAKRYS 103
|||||
Db 181 ELAKRYS 187

RESULT 41

D75161

3-dehydroquinate dehydratase (arod) PAB0299 - Pyrococcus abyssi (strain Orsay)

C:Species: Pyrococcus abyssi

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000

C:Accession: D75161

R:anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s

A:Reference number: A75001

A:Accession: D75161

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-215 <KAW>

A:Cross-references: GB:AJ248284; GB:AL096836; NID:g5457730; PIDN:CAB49371.1; PID:g545

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: arod; PAB0299

C:Superfamily: 3-dehydroquinatase dehydratase; 3-dehydroquinatase dehydratase homology

F;2-201/Domain: 3-dehydroquinatase dehydratase homology <QD>

Query Match 1.8%; Score 7; DB 2; Length 215;

Best Local Similarity 100.0%; Pred. No. 64;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 324 AIEKIKS 330

|||||

Db 14 AIEKIKS 20

RESULT 42

C84786

hypothetical protein At2g36920 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: C84786

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487

A:Accession: C84786

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-216 <STO>

A:Cross-references: GB:AE002093; NID:g4883608; PIDN:AAD31577.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g36920

A:Map position: 2

Query Match 1.8%; Score 7; DB 2; Length 216;

Best Local Similarity 100.0%; Pred. No. 64;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 255 GSIESIL 261

|||||

Db 142 GSIESIL 148

RESULT 43

A45929

spectrin beta chain - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 17-Mar-1999
 C:Accession: A45929
 R:Cioe, L.; Laurila, P.; Meo, P.; Krebs, K.; Goodman, S.; Curtis, P.J.
 Blood 70, 915-920, 1987
 A:Title: Cloning and nucleotide sequence of a mouse erythrocyte beta-spectrin cDNA.
 A:Reference number: A45929; MUID:88001019
 A:Accession: A45929
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-237 <C10>
 A:Cross-references: GB:M18641
 C:Superfamily: spectrin beta chain; alpha-actinin actin-binding domain homology; spectrin
 C:Keywords: actin binding

Query Match 1.8%; Score 7; DB 2; Length 237;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 175 AVASEDM 181
 |||||
 Db 21 AVASEDM 27

RESULT 44

F81363
 Probable glutamine transport ATP-binding protein Cj0902 [imported] - Campylobacter jejuni
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 21-Jul-2000
 C:Accession: F81363
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillin
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre
 Nature 403, 665-668, 2000
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
 A:Reference number: A81250; MUID:20150912
 A:Accession: F81363
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-242 <PAR>
 A:Cross-references: GB:AL139076; GB:AL111168; MID:g6968128; PIDN:CAB73160.1; PID:g696834
 A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: glnQ; Cj0902
 C:Superfamily: inner membrane protein malk; ATP-binding cassette homology

Query Match 1.8%; Score 7; DB 2; Length 242;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 KIAVDAS 36
 |||||
 Db 213 KIAVDAS 219

RESULT 45

F91253
 Probable DNA modification protein [imported] - Escherichia coli (strain O157:H7, substra
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: F91253
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: F91253
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-245 <RAY>
 A:Cross-references: GB:BA000007; PIDN:BA838421.1; PID:g13364474; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:

A:Gene: ECS4998
 C:Superfamily: phage Mu gene mom protein
 Query Match 1.8%; Score 7; DB 2; Length 245;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 247 ALKLIRQ 253
 |||||
 Db 122 ALKLIRQ 128
 Search completed: November 5, 2002, 16:26:04
 Job time : 50 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 13:52:11 ; Search time 25 seconds
(without alignments)
586.988 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 379

Sequence: 1 MGKGLTKLLADNAPKAMKE.....SDKTSKAAANKTKAGGKKK 379

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	13	3.4	378	1	FEN1_MOUSE	P39749	mus musculus
2	13	3.4	380	1	FEN1_HUMAN	P39748	homo sapien
3	11	2.9	380	1	RAD2_SCHPO	P39750	saccharomyc
4	9	2.4	382	1	RAZ7_YEAST	P26793	saccharomyc
5	8	2.1	172	1	GLNA_BOVIN	P15103	bos taurus
6	8	2.1	352	1	RECA_BORPE	P17740	bordetella
7	8	2.1	369	1	GLNA_YEAST	P32288	saccharomyc
8	8	2.1	373	1	GLNA_CHICK	P16580	gallus gall
9	8	2.1	373	1	GLNA_CRILLO	P04773	cricketulus
10	8	2.1	373	1	GLNA_HUMAN	P15104	homo sapien
11	8	2.1	373	1	GLNA_PIG	P48410	sus scrofa
12	8	2.1	559	1	CAFB_HUMAN	Q13112	homo sapien
13	8	2.1	572	1	CAFB_MOUSE	Q940n7	mus musculus
14	7	1.8	80	1	R37A_THEAC	Q9hip0	thermoplasm
15	7	1.8	100	1	REGN_BPP22	P04891	bacterioph
16	7	1.8	147	1	FLAW_DESGI	Q01096	desulfovibr
17	7	1.8	152	1	RL30_METTH	Q26132	methanobact
18	7	1.8	215	1	AROD_PVRAB	Q9v1h8	pyrococcus
19	7	1.8	257	1	GHSR_MOUSE	Q99p50	mus musculus
20	7	1.8	259	1	CTRI_ANOGA	Q27289	anopheles g
21	7	1.8	261	1	SP3J_BACSU	Q01625	bacillus su
22	7	1.8	275	1	NUSE_SYNY3	P74395	synecocyst
23	7	1.8	287	1	SMN_BOVIN	O18870	bos taurus
24	7	1.8	304	1	Y191_METTH	O26293	methanobact
25	7	1.8	314	1	HRPR_PSESY	P37930	pseudomonas
26	7	1.8	325	1	SYFA_THEMEA	Q9wz88	thermotoga
27	7	1.8	364	1	GHSR_RAT	Q08725	rattus norv
28	7	1.8	377	1	YC81_MYCPN	P75496	mycoplasma
29	7	1.8	399	1	GLN1_DROME	P20477	drosophila
30	7	1.8	411	1	YPEB_SYNP2	P42784	synecococc
31	7	1.8	413	1	DCHS_LYCES	P54772	lycopersico
32	7	1.8	449	1	IFR1_MOUSE	P19182	mus musculus
33	7	1.8	449	1	IFR1_RAT	P20695	rattus norv

34	7	1.8	459	1	COAT_FLDV	P22176	fish lympho
35	7	1.8	460	1	YTM1_YEAST	Q12024	saccharomyc
36	7	1.8	472	1	C13B_MYCTU	O53765	mycobacteri
37	7	1.8	474	1	YU88_MYCTU	O53305	mycobacteri
38	7	1.8	488	1	PE24_HUMAN	P35408	homo sapien
39	7	1.8	488	1	PE24_RABIT	O28691	oryctolagus
40	7	1.8	488	1	PE24_RAT	P43114	rattus norv
41	7	1.8	491	1	A2AP_HUMAN	P08697	homo sapien
42	7	1.8	492	1	GTRL_MOUSE	P17809	mus musculus
43	7	1.8	506	1	YFCC_ECOLI	P32663	escherichia
44	7	1.8	507	1	RECN_CAMJE	Q9phm5	campylobact
45	7	1.8	513	1	PE24_MOUSE	P32240	mus musculus

ALIGNMENTS

```
RESULT 1
FEN1_MOUSE
ID FEN1_MOUSE STANDARD; PRT; 378 AA.
AC P39749;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE FLAP endonuclease-1.
GN FEN1 OR FEN-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 260-275 AND 331-352.
RC STRAIN=BALE/C;
RX MEDLINE=95011546; PubMed=7926735;
RA Harrington J.J., Lieber M.R.;
RT "Functional domains within FEN-1 and RAD2 define a family of
RT structure-specific endonucleases: implications for nucleotide excision
RT repair.";
RL Genes Dev. 8:1344-1355(1994).
CC -!- FUNCTION: ENDONUCLEASE THAT CLEAVE 5'FLAP STRUCTURE AND FAILS
CC TO CLEAVE OTHER DNA STRUCTURES, INCLUDING 3'FLAPS AND SINGLE
CC STRANDED DNA.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L26320; AAC37664.1;
CC HSSP; O58839; 1A76.
CC MGD; MGI:102779; Fen1.
CC InterPro; IPR002421; 5_3_exonuclease.
CC InterPro; IPR000513; Exo_N1.
CC InterPro; IPR003584; HHH_2.
CC Pfam; PF01367; 5_3_exonuclease; 1.
CC Pfam; PF00867; XPG_I; 1.
CC Pfam; PF00752; XPG_N; 1.
CC PRINTS; PR00853; XPGRADSUPER.
CC SMART; SM00279; Hh2; 1.
CC SMART; SM00484; XPG1; 1.
CC SMART; SM00485; XPGN; 1.
CC PROSITE; PS00841; XPG_1; 1.
CC PROSITE; PS00842; XPG_2; 1.
CC Hydrolase; Nuclease; Endonuclease; Nuclear protein.
CC DOMAIN 1 102
CC DOMAIN 120 251
CC I-DOMAIN.
CC FT
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X77041; CAB36991.1; -;
 DR EMBL; Z99167; CAB16282.1; -;
 DR HSSP; Q58839; 1A76
 DR InterPro; IPR000513; Exo_N_I.
 DR InterPro; IPR003584; HHL2.
 DR Pfam; PF00867; XPG_I; 1.
 DR Pfam; PF00752; XPG_N; 1.
 DR PRINTS; PR00853; XPGRADSUPER.
 DR SMART; SM00279; HhH2; 1.
 DR SMART; SM00484; XPGI; 1.
 DR SMART; SM00485; XPGN; 1.
 DR PROSITE; PS00841; XPG_1; 1.
 DR PROSITE; PS00842; XPG_2; 1.
 KW DNA repair; Hydrolase; Nuclease; Endonuclease; Nuclear protein.
 FT DOMAIN 1 105 N-
 FT DOMAIN 123 254 I-
 SQ SEQUENCE 380 AA; 42866 MW; 3604065D9B934CBF CRC64;

Query Match 2.9%; Score 11; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 0.0036;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 KRTKVTQHN 139
 DB 129 KRTKVTQHN 139

RESULT 4

ID RAD2_YEAST STANDARD; PRT; 382 AA.
 AC P26793;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Structure specific endonuclease RAD27.
 GN RAD27 OR RTH1 OR YKL13C OR YKL510
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92221689; PubMed=1561835;
 RA Jacquier A., Legrain P., Dujon B.;
 RT "Sequence of a 10.7 kb segment of yeast chromosome XI identifies the
 RT APN1 and the BAF1 loci and reveals one tRNA gene and several new open
 RT reading frames including homologs to RAD2 and kinases.";
 RL Yeast 8:121-132(1992).
 RN [2]
 RP CHARACTERIZATION.
 RC STRAIN=BALE/C;
 RX MEDLINE=95011546; PubMed=7926735;
 RA Harrington J.J., Lieber M.R.;
 RT "Functional domains within FEN-1 and RAD2 define a family of
 RT structure-specific endonucleases: implications for nucleotide
 RT excision repair.";
 RL Genes Dev. 8:1344-1355(1994).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=95113773; PubMed=7814325;
 RA Reagan M.S., Pittenger C., Siede W., Friedberg E.C.;
 RT "Characterization of a mutant strain of Saccharomyces cerevisiae with
 RT a deletion of the RAD27 gene, a structural homolog of the RAD2
 RT nucleotide excision repair gene.";
 RL J. Bacteriol. 177:364-371(1995).
 CC -(-) FUNCTION: SINGLE-STRANDED DNA ENDONUCLEASE (PROBABLE).

CC -(-) SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -(-) SIMILARITY: BELONGS TO THE XPG/RAD2 ENDONUCLEASE FAMILY. FEN1
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; S94804; AAB21998.1; -;
 DR EMBL; Z28113; CAAB1953.1; -;
 DR PIR; S22267; S22267.
 DR HSSP; Q58839; 1A76.
 DR SGD; S0001596; RAD27.
 DR InterPro; IPR002421; 5_3_exonuclease.
 DR InterPro; IPR000513; Exo_N_I.
 DR InterPro; IPR003584; HHL2.
 DR InterPro; IPR001532; XPG_I.
 DR Pfam; PF01367; 5_3_exonuclease; 1.
 DR Pfam; PF00867; XPG_I; 1.
 DR Pfam; PF00752; XPG_N; 1.
 DR PRINTS; PR00853; XPGRADSUPER.
 DR SMART; SM00279; HhH2; 1.
 DR SMART; SM00484; XPGI; 1.
 DR SMART; SM00485; XPGN; 1.
 DR PROSITE; PS00841; XPG_1; 1.
 DR PROSITE; PS00842; XPG_2; 1.
 KW DNA repair; Hydrolase; Nuclease; Endonuclease; Nuclear protein.
 FT DOMAIN 1 105 N-
 FT DOMAIN 120 251 I-
 SQ SEQUENCE 382 AA; 43279 MW; 1F54B08720121C8C CRC64;

Query Match 2.4%; Score 9; DB 1; Length 382;
 Best Local Similarity 100.0%; Pred. No. 0.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 YVFDGKPPD 92
 DB 84 YVFDGKPPD 92

RESULT 5

ID GLNA_BOVIN STANDARD; PRT; 172 AA.
 AC P15103; O02850;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Glutamine synthetase (EC 6.3.1.2) (Glutamate-ammonia ligase)
 DE (Fragment).
 GN GLUL.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE OF 1-66 FROM N.A.
 RA Masabanda J., Wigger G., Eggen A., Stranzinger G., Fries R.;
 RT "The bovine glutamine synthase gene (GLUL) maps to 10q33 and a
 RT pseudogene (GLULP) to 16q21.";
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 66-172 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=88124783; PubMed=2893372;
 RA Smith D.D. Jr., Campbell J.W.;
 RT "Distribution of glutamine synthetase and carbamoyl-phosphate
 RT synthetase I in vertebrate liver.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:160-164(1988).


```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M65157; AAA34644.1; ALT_SEQ.
DR EMBL; 268111; CAA92141.1; -.
DR EMBL; 271255; CAA94985.1; -.
DR EMBL; 249274; CAA89289.1; -.
DR PIR; A42278; A42278.
DR HSSP; P06201; ILGR.
DR SWISS-2DPAGE; P32288; YEAST.
DR YEPD; 5410; -.
DR SGD; S0006239; GLN1.
DR InterPro; IPR001691; GLN_synth.
DR Pfam; PF00120; gln-synt; 1.
DR PROSITE; PS00180; GLNA_1; 1.
DR PROSITE; PS00181; GLNA_ATP; 1.
KW Ligase; Acetylation.
FT INIT_MET 0 0
FT MOD_RES 1 1 ACETYLATION.
FT CONFLICT 164 164 G -> D (IN REF. 3).
FT CONFLICT 171 171 M -> V (IN REF. 3).
FT CONFLICT 250 250 T -> A (IN REF. 2).
FT CONFLICT 263 263 M -> T (IN REF. 2).
SQ SEQUENCE 369 AA; 41635 MW; 4A1E71B684D4BC99 CRC64;

Query Match 2.1%; Score 8; DB 1; Length 369;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 ATEKLSKR 130
Db 269 ATEKLSKR 276

RESULT 8
GLNA_CHICK
ID GLNA_CHICK STANDARD; PRT; 373 AA.
AC P16580;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase).
GN GLUL OR GLNS.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RE SEQUENCE FROM N.A.
RX MEDLINE=90034178; PubMed=2572512;
RA Pu H., Young A.P.;
RT "The structure of the chicken glutamine synthetase-encoding gene.";
RL Gene 81:169-175(1989).
RN [2]
RE SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92408449; PubMed=1356223;
RA Campbell J.W., Smith D.D. Jr.;
RT "Metabolic compartmentation of vertebrate glutamine synthetase:
RT putative mitochondrial targeting signal in avian liver glutamine
RT synthetase.";
RL Mol. Biol. Evol. 9:787-805(1992).
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC L-glutamine.
CC -!- SUBUNIT: HOMOOCTAMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

```

```

CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S45408; RAC69361.1; -.
DR EMBL; M29076; AAA48783.1; -.
DR PIR; J00025; AJCHQ.
DR PIR; A44095; A44095.
DR HSSP; P06201; 2LGS.
DR InterPro; IPR001691; GLN_synth.
DR Pfam; PF00120; gln-synt; 1.
DR PROSITE; PS00180; GLNA_1; 1.
DR PROSITE; PS00181; GLNA_ATP; 1.
KW Ligase.
SQ SEQUENCE 373 AA; 42146 MW; 505254A25E8733DB CRC64;

Query Match 2.1%; Score 8; DB 1; Length 373;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 ATEKLSKR 130
Db 273 ATEKLSKR 280

RESULT 9
GLNA_CRILLO
ID GLNA_CRILLO STANDARD; PRT; 373 AA.
AC P04773;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase).
GN GLUL.
OS Cricetulus longicaudatus (Long-tailed hamster) (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10030;
RN [1]
RE SEQUENCE FROM N.A.
RX MEDLINE=86120385; PubMed=2868445;
RA Hayward B.E., Hussain A., Wilson R.H., Lyons A., Woodcock V.,
RA McIntosh B., Harris T.J.R.;
RT "The cloning and nucleotide sequence of cDNA for an amplified
RT glutamine synthetase gene from the Chinese hamster.";
RL Nucleic Acids Res. 14:999-1008(1986).
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC L-glutamine.
CC -!- SUBUNIT: HOMOOCTAMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03495; CAA27211.1; -.
DR EMBL; A12480; CAA01036.1; -.
DR PIR; A23524; AJHHQ.
DR InterPro; IPR001691; GLN_synth.
DR Pfam; PF00120; gln-synt; 1.
DR PROSITE; PS00180; GLNA_1; 1.

```

```

DR PROSITE: PS00181; GLNA_ATP; 1.
KW Ligase.
SQ SEQUENCE 373 AA; 42137 MW; DE8AE6275D2EBC2E CRC64;

Query Match
Best Local Similarity 2.1%; Score 8; DB 1; Length 373;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
Db 273 AIEKLSKR 280

RESULT 10
GLNA_HUMAN
ID GLNA_HUMAN STANDARD; PRT; 373 AA.
AC P15104;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase).
GN GLUL OR GLNS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=87316885; PubMed=2888076;
RA Gibbs C.S., Campbell K.E., Wilson R.H.;
RT "Sequence of a human glutamine synthetase cDNA.";
RL Nucleic Acids Res. 15:6293-6293(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92031701; PubMed=1681907;
RA van den Hoff M.J.B., Geerts W.J.C., Das A.T., Moorman A.F.M.,
RA Lamers W.H.;
RT "cDNA sequence of the long mRNA for human glutamine synthetase.";
RL Biochim. Biophys. Acta 1090:249-251(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94229472; PubMed=7909780;
RA Christa L., Simon M.T., Flinois J.P., Gebhardt R., Brechot C.,
RA Lasserre C.;
RT "Overexpression of glutamine synthetase in human primary liver
RT cancer.";
RL Gastroenterology 106:1312-1320(1994).
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC L-glutamine.
CC -!- SUBUNIT: HOMOOCTAMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00387; CAA68457.1; -
DR EMBL: X59834; CAA2495.1; -
DR EMBL: S70290; AAB30693.1; -
DR PIR: S00673; AJHUQ.
DR PIR: S15485; S15485.
DR HSSP: P06201; ILGR.
DR MIM: 138290; -
DR InterPro: IPR001691; GLN_synth.

Query Match
Best Local Similarity 2.1%; Score 8; DB 1; Length 373;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
Db 273 AIEKLSKR 280

RESULT 12
CAFE_HUMAN
ID CAFE_HUMAN STANDARD; PRT; 559 AA.
AC Q13112; Q99548;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)

```

```

DR PFAM: PF00120; qln-synt; 1.
DR PROSITE: PS00180; GLNA_1; 1.
DR PROSITE: PS00181; GLNA_ATP; 1.
KW Ligase.
FT CONFLICT 314 314 A -> G (IN REF. 1).
FT CONFLICT 322 323 SI -> RL (IN REF. 2).
FT CONFLICT 347 347 D -> E (IN REF. 2).
SQ SEQUENCE 373 AA; 42064 MW; 45390C100924FAF3 CRC64;

Query Match
Best Local Similarity 2.1%; Score 8; DB 1; Length 373;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
Db 273 AIEKLSKR 280

RESULT 11
GLNA_PIG
ID GLNA_PIG STANDARD; PRT; 373 AA.
AC P46410;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase).
GN GLUL OR GLNS.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Johnstone R.W., Loveland B.E.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC L-glutamine.
CC -!- SUBUNIT: HOMOOCTAMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z29636; CAA82747.1; -
DR InterPro: IPR001691; GLN_synth.
DR PIR: PF00120; qln-synt; 1.
DR PROSITE: PS00180; GLNA_1; 1.
DR PROSITE: PS00181; GLNA_ATP; 1.
KW Ligase.
SQ SEQUENCE 373 AA; 42030 MW; FB0FD3414ACA7C55 CRC64;

Query Match
Best Local Similarity 2.1%; Score 8; DB 1; Length 373;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
Db 273 AIEKLSKR 280

RESULT 12
CAFE_HUMAN
ID CAFE_HUMAN STANDARD; PRT; 559 AA.
AC Q13112; Q99548;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)

```

DE Chromatin assembly factor 1 subunit B (CAF-1 subunit B) (Chromatin
DE assembly factor I p60 subunit) (CAF-I 60 kDa subunit) (CAF-Ip60) (M-
DE phase phosphoprotein 7).
GN CHAF1B OR CAF1A OR CAF1P60 OR MPP7 OR MPHOSPH7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=9532396; PubMed=7600578;
RA Kaufman P.D., Kobayashi R., Kessler N., Stillman B.;
RT "The p150 and p60 subunits of chromatin assembly factor 1: a
RT molecular link between newly synthesized histones and DNA
RT replication.";
RL Cell 81:1105-1114(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Shibuya K., Kudoh J., Minoshima S., Kawasaki K., Nakatoh E.,
RA Shintani A., Asakawa S., Shimizu N.;
RT "Genomic sequencing of 1.2-Mb region on human chromosome 21q22.2.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=2028799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
RA Ohki M., Takagi T., Sakaki Y., Tauden S., Blechschmidt K., Polley A.,
RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Zimmermann W.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordstok G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ransner J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA Wehrmeyer S., Borzlyn K., Gardino K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
RN [4]
RP SEQUENCE OF 474-559 FROM N.A., AND PHOSPHORYLATION.
RX TISSUE=Lymphoblast;
RC MEDLINE=97039687; PubMed=8885239;
RA Matsumoto-Taniura N., Pirollet F., Monroe R., Gerace L.,
RA Westendorp J.M.;
RT "Identification of novel M phase phosphoproteins by expression
RT cloning.";
RL Mol. Biol. Cell 7:1455-1469(1996).
RN [5]
RP FUNCTION.
RX MEDLINE=99030822; PubMed=9813080;
RA Martini E., Roche D.M., Marheineke K., Verreault A., Almouzni G.;
RT "Recruitment of phosphorylated chromatin assembly factor 1 to
RT chromatin after UV irradiation of human cells.";
RL J. Cell Biol. 143:563-575(1998).
RN [6]
RP REVIEW.
RX MEDLINE=20351441; PubMed=10893180;
RA Ridgway P., Almouzni G.;
RT "CAF-1 and the inheritance of chromatin states: at the crossroads of
RT DNA replication and repair.";
RL J. Cell Sci. 113:2647-2659(2000).
RN [7]
RP SUBCELLULAR LOCATION, AND PHOSPHORYLATION.
RX MEDLINE=98279052; PubMed=9614144;
RA Marheineke K., Krude T.;
RT "Nucleosome assembly activity and intracellular localization of human
RT CAF-1 changes during the cell division cycle.";
RL J. Biol. Chem. 273:15279-15286(1998).
CC -!- FUNCTION: COMPLEX THAT IS THOUGHT TO MEDIATE CHROMATIN ASSEMBLY IN
CC DNA REPLICATION AND DNA REPAIR. ASSEMBLES HISTONE OCTAMERS ONTO
CC REPLICATING DNA IN VITRO. CAF-1 PERFORMS THE FIRST STEP OF THE
CC NUCLEOSOME ASSEMBLY PROCESS, BRINGING NEWLY SYNTHESIZED HISTONES

CC H3 AND H4 TO REPLICATING DNA; HISTONES H2A/H2B CAN BIND TO THIS
CC CHROMATIN PRECURSOR SUBSEQUENT TO DNA REPLICATION TO COMPLETE THE
CC HISTONE OCTAMER.
CC -!- SUBUNIT: CAF-1 IS COMPOSED OF THREE SUBUNITS, P48, P60 AND P150.
CC ONLY MINOR AMOUNTS OF P48 ARE COMPLEXED WITH P60 AND P150 IN G1
CC PHASE. IN G2 AND S PHASE ALSO MONOMERIC P60 IS DETECTED.
CC -!- SUBCELLULAR LOCATION: NUCLEAR; DNA REPLICATION FOCI. CYTOPLASMIC
CC IN M PHASE.
CC -!- DEVELOPMENTAL STAGE: ACTIVE COMPLEX IS FOUND IN G1, S AND G2
CC PHASES.
CC -!- PTM: DIFFERENTIALLY PHOSPHORYLATED DURING CELL CYCLE. DURING
CC MITOSIS THE P60 SUBUNIT OF INACTIVE CAF-1 IS HYPERPHOSPHORYLATED
CC AND DISPLACED INTO THE CYTOSOL. PROGRESSIVELY DEPHOSPHORYLATED FROM
CC G1 TO S AND G2 PHASE. PHOSPHORYLATED P60 IS RECRUITED TO CHROMATIN
CC UNDERGOING DNA REPAIR AFTER UV IRRADIATION IN G1, S OR G2 PHASES.
CC -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: BELONGS TO THE HIR1 FAMILY OF WD-REPEAT
CC PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U20980; AAA76737.1; -;
CC EMBL: AF000694; BAA89426.1; -;
CC EMBL: AP001725; BAA95549.1; -;
CC EMBL: X98262; CAA66915.1; -;
CC MIM: 601245; -;
CC InterPro: IPR001680; WD40.
CC Pfam: PF00400; WD40. 5.
CC PRINTS: P003320; GPROTEINBRPT.
CC SMART: SM00320; WD40; 5.
CC PROSITE: PS00678; WD_REPEATS_1; 1.
CC PROSITE: PS50082; WD_REPEATS_2; 3.
CC PROSITE: PS50294; WD_REPEATS_REGION; 1.
CC DNA replication; DNA repair; Cell cycle; Nuclear protein;
CC Phosphorylation; Repeat; WD repeat.
CC REPEAT 11. 54
CC REPEAT 64. 103 WD 2.
CC REPEAT 127. 166 WD 3.
CC REPEAT 169. 208 WD 4.
CC REPEAT 228. 279 WD 5.
CC REPEAT 299. 340 WD 6.
CC REPEAT 344. 385 WD 7.
CC MOD_RES 495. 495 PHOSPHORYLATION (POTENTIAL).
CC MOD_RES 503. 503 PHOSPHORYLATION (POTENTIAL).
CC MOD_RES 509. 509 PHOSPHORYLATION (POTENTIAL).
CC MOD_RES 521. 521 PHOSPHORYLATION (POTENTIAL).
CC MOD_RES 531. 531 PHOSPHORYLATION (POTENTIAL).
CC CONFLICT 494. 494 K -> N (IN REF. 4).
CC SEQUENCE 559 AA; 61492 MW; AD1846CG81B8DC9F CRC64;
Query Match 2.1%; Score 8; DB 1; Length 559;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 173 VFAVASED 180
Db 318 VFAVASED 325
|||||||
RESULT 13
ID CAFB_MOUSE STANDARD; PRT; 572 AA.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chromatin assembly factor 1 subunit B (CAF-1 subunit B) (Chromatin
DE assembly factor I p60 subunit) (CAF-I 60 kDa subunit) (CAF-Ip60).
DE

GN CHAF1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RA "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
CC -!- FUNCTION: COMPLEX THAT IS THOUGHT TO MEDIATE CHROMATIN ASSEMBLY IN
CC DNA REPLICATION AND DNA REPAIR. ASSEMBLES HISTONE OCTAMERS ONTO
CC REPLICATING DNA IN VITRO. CAF-1 PERFORMS THE FIRST STEP OF THE
CC NUCLEOSOME ASSEMBLY PROCESS, BRINGING NEWLY SYNTHESIZED HISTONES
CC H3 AND H4 TO REPLICATING DNA; HISTONES H2A/H2B CAN BIND TO THIS
CC CHROMATIN PRECURSOR SUBSEQUENT TO DNA REPLICATION TO COMPLETE THE
CC HISTONE OCTAMER (BY SIMILARITY).
CC -!- SUBUNIT: CAF-1 IS COMPOSED OF THREE SUBUNITS, P48, P60 AND P150.
CC -!- SUBCELLULAR LOCATION: NUCLEAR; DNA REPLICATION FOCI (BY
CC similarity).
CC -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: BELONGS TO THE HIRL FAMILY OF WD-REPEAT PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AK011243; BAB27490.1;
CC MGD: MGI:1314881; Chaf1b.
CC InterPro: IPR001632; Gprotein_B.
CC InterPro: IPR001680; WD40.
CC Pfam: PF00400; WD40; 5.
CC PRINTS: PR00319; GPROTEINB.
CC PRINTS: PR00320; GPROTEINRPT.
CC SMART: SM00320; WD40; 6.
CC PROSITE: PS00678; WD_REPEATS_1; 1.
CC PROSITE: PS00082; WD_REPEATS_2; 3.
CC PROSITE: PS50294; WD_REPEATS_REGION; 1.
CC DNA replication; DNA repair; Cell cycle; Nuclear protein;
CC Phosphorylation; Repeat; WD repeat.
FT REPEAT 11 54
FT REPEAT 64 103 WD 1.
FT REPEAT 127 166 WD 2.
FT REPEAT 169 208 WD 3.
FT REPEAT 228 279 WD 4.
FT REPEAT 301 347 WD 5.
FT REPEAT 351 392 WD 6.
FT REPEAT 502 502 WD 7.
FT MOD_RES 502 502 PHOSPHORYLATION (POTENTIAL).
FT MOD_RES 510 510 PHOSPHORYLATION (POTENTIAL).
SQ SEQUENCE 572 AA; 63132 MW; 03C2BC263591FF72 CRC64;

Query Match 2.1%; Score 8; DB 1; Length 572;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 VFAVASED 180
DB 325 VFAVASED 332
ID R37A_THEAC STANDARD; PRT; 80 AA.
AC Q9HIP0:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L37AE.
GN RPL37AE OR TAI295.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;
OC Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Meves H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.,
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum."
RL Nature 407:508-513(2000).
CC -!- SIMILARITY: BELONGS TO THE L37AE FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AL445067; CAC12417.1;
CC InterPro: IPR002674; Ribosomal_L37ae.
CC Pfam: PF01780; Ribosomal_L37ae; 1.
CC ProDom: PD006591; Ribosomal_L37ae; 1.
CC Ribosomal protein; RNA-binding; Zinc-finger; Metal-binding;
CC Complete proteome.
FT ZN.FING 39 60 C4-TYPE.
SQ SEQUENCE 80 AA; 9120 MW; 6F8915ECA5866850 CRC64;

Query Match 1.8%; Score 7; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 SKRTVKV 134
DB 2 SKRTVKV 8
ID REGN_BPP22 STANDARD; PRT; 100 AA.
AC P04891;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable regulatory protein N.
GN N OR 24.
OS Bacteriophage P22.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae.
OX NCBI_TaxID=10754;
RN [1]

```
RP SEQUENCE FROM N.A.
RX MEDLINE=85160834; PubMed=3157001;
RA Franklin N.C.;
RT "Conservation of genome form but not sequence in the transcription
RT antitermination determinants of bacteriophages lambda, phi 21 and
RT P22."
RL J. Mol. Biol. 181:75-84(1985).
[2]
RN REVISIONS, SEQUENCE OF 69-100 FROM N.A.
RP MEDLINE=93328675; PubMed=8335629;
RA Ranade K., Poteete A.R.;
RT "Superinfection exclusion (sieB) genes of bacteriophages P22 and
RT lambda."
RL J. Bacteriol. 175:4712-4718(1993).
[3]
RN SEQUENCE FROM N.A.
RP Kropinski A.M.B., VanderByl C.S.;
RT "The completed sequence of genome of Salmonella phage P22."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[4]
RN STRUCTURE BY NMR OF 14-32.
RX MEDLINE=98162556; PubMed=9501914;
RA Cai Z., Gorin A., Frederick R., Ye X., Hu W., Majumdar A.,
RA Kettani A., Patel D.J.;
RT "Solution structure of P22 transcriptional antitermination N peptide-
RT boxB RNA complex."
RL Nat. Struct. Biol. 5:203-212(1998).
CC -!- FUNCTION: N PROTEIN REGULATES THE TRANSITION FROM THE EARLY TO
CC THE MIDDLE STAGE OF LYtic DEVELOPMENT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X02141; CAA26079.1; ALT_SEQ.
DR EMBL; L18800; -; NOT_ANNOTATED_CDS.
DR EMBL; AF217253; AAF75023.1; -.
DR PIR; S10069; S10069.
DR PDB; 1A4T; 29-APR-98.
KW Transcription regulation; Transcription termination; 3D-structure.
SQ SEQUENCE 100 AA; 10970 MW; E7528F49C4E1A3BB CRC64;

Query Match 1.8%; Score 7; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 DRVTKAI 325
DB 58 DRVTKAI 64

RESULT 16
FLAV_DESGI
ID FLAV_DESGI STANDARD; PRT; 147 AA.
AC Q01096;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Flavodoxin.
OS Desulfovibrio gigas.
OC Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.
OX NCBI_TaxID=879;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 29494 / DSM 496;
RX MEDLINE=92329549; PubMed=1627649;
RA Helms L.R., Swenson R.P.;
RT "The primary structures of the flavodoxins from two strains of
RT Desulfovibrio gigas. Cloning and nucleotide sequence of the
```

```
RT structural genes."
RL Biochim. Biophys. Acta 1131:325-328(1992).
CC -!- FUNCTION: LOW-POTENTIAL ELECTRON DONOR TO A NUMBER OF REDOX
CC ENZYMES.
CC -!- COFACTOR: FMN.
CC
CC -!- SIMILARITY: BELONGS TO THE FLAVODOXIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X64765; CAA46012.1; -.
DR PIR; S24310; S24310.
DR HSSP; P00323; 2FX2.
DR InterPro; IPR001226; Flavodoxin.
DR Pfam; PF00258; flavodoxin; 1.
DR PROSITE; PS00201; FLAVODOXIN; 1.
KW Electron transport; Flavoprotein; FMN.
SQ SEQUENCE 147 AA; 15186 MW; 5779A72DD1395635 CRC64;

Query Match 1.8%; Score 7; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 210 VAKVLEE 216
DB 22 VAKVLEE 28

RESULT 17
RL30_METH
ID RL30_METH STANDARD; PRT; 152 AA.
AC Q26132;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L30P.
GN RPL30P OR MTH24.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=DELTA H.
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics."
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- SIMILARITY: BELONGS TO THE L30P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AE000796; AAB84533.1; -.
DR InterPro; IPR000517; Ribosomal_L30.
DR Pfam; PF00327; Ribosomal_L30; 1.
```

```
DR PROSITE; PS00634; RIBOSOMAL_L30; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 152 AA; 17109 MW; 4E10395EA96D5F44 CRC64;

Query Match 1.8%; Score 7; DB 1; Length 152;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 77 EAGIKPV 83
    |||||
DB 108 EAGIKPV 114

RESULT 18
AROD_PYRAB STANDARD; PRT; 215 AA.
AC Q9V1H8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-dehydroquininate dehydratase (EC 4.2.1.10) (3-dehydroquinase)
DE (Type I DHQase).
GN AROD OR PAB0299.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Hellig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
    structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: 3-dehydroquininate = 3-dehydroshikimate + H(2)O.
CC -!- PATHWAY: THIRD STEP IN THE BIOSYNTHESIS OF CHOLISMATE WITHIN
    THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC -!- SIMILARITY: BELONGS TO THE TYPE-I 3-DEHYDROQUINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
    the European Bioinformatics Institute. There are no restrictions on its
    use by non-profit institutions as long as its content is in no way
    modified and this statement is not removed. Usage by and for commercial
    entities requires a license agreement (See http://www.isb-sib.ch/announce/
    or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ248284; CAB49371.1; -.
DR InterPro; IPR001381; DHQinase_1.
DR Pfam; PF01487; DHQinase_1; 1.
DR PROSITE; PS01028; DEHYDROQUINASE_I; FALSE_NEG.
KW Aromatic amino acid biosynthesis; Lyase; Complete proteome.
FT ACT_SITE 112 112 BY SIMILARITY.
FT ACT_SITE 139 139 FORMS A SCHIFF-BASE INTERMEDIATE
    (BY SIMILARITY).
SQ SEQUENCE 215 AA; 24510 MW; 02CC027484DB85D1 CRC64;

Query Match 1.8%; Score 7; DB 1; Length 215;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 324 AIEIKTS 330
    |||||
DB 14 AIEIKTS 20

RESULT 19
GHSR_MOUSE STANDARD; PRT; 257 AA.
AC Q99P50;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Growth hormone secretagogue receptor type 1 (GHS-R) (GH-releasing

peptide receptor) (GHRP) (Ghrelin receptor) (Fragment).
GN GHSR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-183 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hypothalamus;
RA Kacsob B.;
RT "Cloning of mouse ghrelin/growth hormone secretagogue receptor cDNA by
    rapid amplification of cDNA ends (RACE).";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 73-257 FROM N.A.
RC STRAIN=129S3/SvImJ;
RA Peng X., Frohman L.A., Kineman R.D.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for ghrelin, coupled to G-alpha-11 proteins.
    Stimulates growth hormone secretion. Binds also other growth
    hormone releasing peptides (GHRP) (e.g. Met-enkephalin and GHRP-6)
    as well as non-peptide, low molecular weight secretagogues (e.g.
    L-692,429, MK-0677, adenosine) (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
    between the Swiss Institute of Bioinformatics and the EMBL outstation -
    the European Bioinformatics Institute. There are no restrictions on its
    use by non-profit institutions as long as its content is in no way
    modified and this statement is not removed. Usage by and for commercial
    entities requires a license agreement (See http://www.isb-sib.ch/announce/
    or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AY056474; AAL13336.1; -.
DR EMBL; AF332997; AAG61141.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 35 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 36 58 1 (POTENTIAL).
FT DOMAIN 59 77 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 78 100 2 (POTENTIAL).
FT DOMAIN 101 125 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 126 148 3 (POTENTIAL).
FT DOMAIN 149 160 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 161 183 4 (POTENTIAL).
FT DOMAIN 184 211 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 212 234 5 (POTENTIAL).
FT DOMAIN 235 257 CYTOPLASMIC (POTENTIAL).
FT DISULFID 115 197 BY SIMILARITY.
FT CARBOHYD 13 13 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 187 187 N-LINKED (GLCNAC. .) (POTENTIAL).
FT NON_TER 257 257
SQ SEQUENCE 257 AA; 28743 MW; D6FA21EAB0E30791 CRC64;

Query Match 1.8%; Score 7; DB 1; Length 257;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 286 EPNVTL 292
    |||||
DB 11 EPNVTL 17

RESULT 20
CTRL_ANOGA STANDARD; PRT; 259 AA.
ID CTRL_ANOGA
AC Q27289;
```



```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chymotrypsin I precursor (EC 3.4.21.1).
GN CHYMI.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SUAKOKO; TISSUE=Midgut;
RA Mueller H.M.;
RL Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SUAKOKO; TISSUE=Midgut;
RA Mueller H.M.; Catteruccia F.; Crisanti A.;
RT "An Anopheles gambiae locus containing the sequences of two closely
RT related chymotrypsin-like proteases induced in the gut following blood
RT meal.";
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Tyr-I-Xaa, Trp-I-Xaa,
CC Phe-I-Xaa, Leu-I-Xaa.
CC -!- TISSUE SPECIFICITY: MIDGUT.
CC -!- DEVELOPMENTAL STAGE: INDUCED IN THE MIDGUT OF FEMALE AFTER BLOOD
CC MEAL.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 218887; CAA79325.1; -
CC EMBL; 232645; CAA83568.1; -
CC HSP; P00766; ICHG.
CC MEROPS: S01.121.1.
CC InterPro: IPR001314; Chymotrypsin.
CC DR IPfam; IPR001254; Trypsin.
CC DR Pfam; PF00089; trypsin; 1.
CC DR PRINTS; PR00722; CHYMOTRYPSIN.
CC DR SMART; SM00020; Tryp_Spc; 1.
CC DR PROSITE; PS02040; TRYPSIN_DOM; 1.
CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
CC KW Hydroxylase; Serine protease; Signal; Zymogen; Digestion;
KW Multigene family.
FT SIGNAL 1 17 POTENTIAL.
FT PROPEP 18 32 ACTIVATION PEPTIDE.
FT CHAIN 33 259 CHYMOTRYPSIN 1.
FT ACT_SITE 74 74 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 119 119 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 59 75 BY SIMILARITY.
FT DISULFID 182 198 BY SIMILARITY.
FT DISULFID 208 232 BY SIMILARITY.
FT SITE 206 206 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 259 AA; 27717 MW; 66A9C0BA9D1E8C68 CRC64;

Query Match 1.88; Score 7; DB 1; Length 259;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 KVFavas 178
Db 4 KVFavas 10

```

```

RESULT 21
SP3J_BACSU STANDARD; PRT; 261 AA.
ID Q01625; O32298;
AC 01-APR-1993 (Rel. 25, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Stage III sporulation protein J precursor.
GN SPOIJJ.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=93139756; PubMed=1487728;
RA Errington J., Appleby L., Daniel R.A., Goodfellow H.,
RA Partridge S.R., Yudkin M.D.;
RT "Structure and function of the spoIJJ gene of Bacillus subtilis: a
RT vegetatively expressed gene that is essential for sigma G activity at
RT an intermediate stage of sporulation.";
RL J. Gen. Microbiol. 138:2609-2618(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / CRK2000;
RX MEDLINE=92204018; PubMed=1552862;
RA Ogasawara N., Yoshikawa H.;
RT "Genes and their organization in the replication origin region of the
RT bacterial chromosome.";
RL Mol. Microbiol. 6:629-634(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogasawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin.";
RL DNA Res. 1:1-14(1994).
CC -!- FUNCTION: ESSENTIAL FOR SIGG ACTIVITY AT AN INTERMEDIATE STAGE OF
CC SPORULATION.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY IN VEGETATIVE CELLS.
CC -!- DEVELOPMENTAL STAGE: SPOIJJ IS REQUIRED ONLY AFTER ENGULFMENT AND
CC TURNED OFF AT THE ONSET OF SPORULATION.
CC -!- SIMILARITY: BELONGS TO THE OXAL / 60 KDA IMP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z14225; CAA78595.1; ALT_INT.
CC EMBL; X62539; CAA44401.1; -
CC EMBL; D26185; BAA05234.1; -
CC EMBL; Z99124; CAB16141.1; -
CC PIR; S18073; S18073.
CC Subtilisin; BG10062; spoIJJ.
CC InterPro: IPR001708; 60kd_inner_MP.
CC Pfam; PF02096; 60KD_IMP_2.
CC PRINTS; PR00701; 60KDINNERMP.
CC PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
CC Sporulation; Membrane; Lipoprotein; Signal; Complete proteome.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 261 STAGE III SPORULATION PROTEIN J.
FT LIPID 23 23 N-ACYL DIGLYCERIDE (POTENTIAL).
SQ SEQUENCE 261 AA; 29521 MW; F21120088CEBEC5 CRC64;

```

```

Query Match          1.8%; Score 7; DB 1; Length 261;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 373 KAGGKKK 379
DB 255 KAGGKKK 261

RESULT 22
NUSB_SVNV3
ID NUSB_SVNV3 STANDARD; PRT; 275 AA.
AC P74395;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE N utilization substance protein B homolog (NusB protein).
GN NUSB OR SLL0271.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA RES. 3:109-136(1996).
CC -!- FUNCTION: INVOLVED IN THE TRANSCRIPTION TERMINATION PROCESS
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE NUSB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90914; BAA18492.1; -
DR HSSP; P95020; 1EYV.
DR InterPro; IPR000139; NusB.
DR Pfam; PF01029; NusB; 1.
DR ProDom; PD005242; NusB; 1.
KW Transcription termination; RNA-binding; Complete proteome.
SQ SEQUENCE 275 AA; 30590 MW; D83B8791A5EA6EB CRC64;

Query Match          1.8%; Score 7; DB 1; Length 275;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 97 ELAKRYS 103
DB 238 ELAKRYS 244

RESULT 23
SMN_BOVIN STANDARD; PRT; 287 AA.
AC O18570; O62700; Q9TS79; O46481;
DT 15-JUL-1998 (Rel. 36, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Survival motor neuron protein.
GN SMN1 OR SMN.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Pietrowski D., Foerster M.;
RT "Complete coding sequence of the bovine smn encoding gene.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 35-123 FROM N.A.
RC STRAIN-BROWN SWISS;
RA Rieder S., Joerg H., Neuschwander S., Meijerink E., Stranzinger G.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 90-171 FROM N.A.
RA Nonnenan D., Shibuya H., Kappes S., Steffen D., Johnson G.S.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 99-141 FROM N.A.
RA Eggen A.;
RT "The bovine survival motor neuron gene (SMN) is located on bovine
RT chromosome 20q12-q13.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE SMN COMPLEX PLAYS AN ESSENTIAL ROLE IN SPLICEOSOMAL
CC SNRP ASSEMBLY IN THE CYTOPLASM, AND IS REQUIRED FOR PRE-MRNA
CC SPLICING IN THE NUCLEUS (BY SIMILARITY).
CC -!- SUBUNIT: FORMS A STABLE HETEROMERIC COMPLEX WITH SURVIVAL OF MOTOR
CC NEURON PROTEIN INTERACTING PROTEIN 1 (SIP1/GEMIN2), GEMIN3 AND
CC GEMIN4. THIS COMPLEX IS THOUGHT TO PLAY AN ESSENTIAL ROLE IN
CC BIOGENESIS OF THE SMALL NUCLEAR RIBONUCLEOPROTEIN PARTICLES
CC (SNRNPs). INTERACTS WITH FIBRILLARIN AND WITH SEVERAL SPLICEOSOMAL
CC SNRP CORE SM PROTEINS, INCLUDING B/B', SM D AND SM E PROTEINS AND
CC WITH ILF3 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC AND NUCLEAR. LOCALIZED IN
CC SUBNUCLEAR STRUCTURES NEXT TO COILED BODIES, CALLED GEMs, WHICH
CC ARE HIGHLY ENRICHED IN SPLICEOSOMAL SNRNPs (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SMN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 TUDOR DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF035322; AAC17995.1; -
DR EMBL; AF035323; AAC63439.1; -
DR EMBL; AF026810; AAB80943.1; -
DR EMBL; AF016590; AAC04667.1; -
DR EMBL; AF034259; AAD01979.1; -
DR InterPro; IPR002999; Tudor.
DR SMART; SM00333; TUDOR; 1.
DR PROSITE; PS50304; TUDOR; 1.
KW mRNA processing; RNA-binding; Nuclear protein.
FT DOMAIN 86 146 TUDOR.
FT DOMAIN 2 7 POLY-GLY.
FT DOMAIN 190 193 POLY-PRO.
FT DOMAIN 212 220 POLY-PRO.
FT DOMAIN 238 242 POLY-PRO.
FT CONFLICT 61 61 K -> E (IN REF. 2).
FT CONFLICT 64 64 G -> A (IN REF. 2).
FT CONFLICT 168 168 MISSING (IN REF. 3).
SQ SEQUENCE 287 AA; 31327 MW; B1B7EFC2E682A78 CRC64;

Query Match          1.8%; Score 7; DB 1; Length 287;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 329 KSAKNKS 335
DB 69 KSAKNKS 75

```

```

RESULT 24
Y191_METTH
ID Y191_METTH STANDARD; PRT; 304 AA.
AC O26293;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamine amidotransferase-like protein WTH191.
GN WTH191.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=3371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwni N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT delTAH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -!- SIMILARITY: THE GATASE DOMAIN BELONGS TO TYPE-2 GLUTAMINE
CC AMIDOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U03853; AAA17651.1; .
CC InterPro: IPR002197; HTH_Fis.
CC InterPro: IPR002078; Sig54_interact.
CC Pfam; PF02954; HTH_8; 1.
CC Pfam; PF00158; Sigma54_activat; 1.
CC PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
CC PROSITE; PS00676; SIGMA54_INTERACT_2; FALSE_NEG.
CC PROSITE; PS00688; SIGMA54_INTERACT_3; FALSE_NEG.
CC PROSITE; PS00443; GATASE_TYPE_II; FALSE_NEG.
CC TRANSFERASE; Glutamine amidotransferase; Complete proteome.
FT INIT_MET 0 0 BY SIMILARITY.
FT ACT_SITE 1 1 GATASE (BY SIMILARITY).
SQ SEQUENCE 304 AA; 33779 MW; 02F652DE0FF6FC89 CRC64;

Query Match 1.8%; Score 7; DB 1; Length 304;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 VFVAVASE 179
DB 272 VFVAVASE 278
|||||||

RESULT 25
HRPR_PSESY
ID HRPR_PSESY STANDARD; PRT; 314 AA.
AC P37930;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pathogenicity locus probable regulatory protein hrpr.
GN HRPR.
OS Pseudomonas syringae (pv. syringae).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=321;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=PSS61;
RX MEDLINE=94148760; PubMed=8106313;
RA Xiao Y., Heu S., Yi J., Lu Y., Hutcheson S.W.;
RT "Identification of a putative alternate sigma factor and
RT characterization of a multicomponent regulatory cascade controlling
RT the expression of Pseudomonas syringae pv. syringae Pss61 hrp and
RT hrpA genes.";
RL J. Bacteriol. 176:1025-1036(1994).
CC -!- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM HRPR/HRPS
CC THAT REGULATES THE ACTIVATION OF THE SIGMA FACTOR HRPL WHICH ARE
CC INVOLVED IN PLANT PATHOGENICITY, HRMA AND AVR GENES. PROBABLY
CC INTERACTS WITH SIGMA-54.
CC -!- SIMILARITY: THE N-TERMINAL REGION CONTAINS A SIGMA-54 FACTOR
CC INTERACTION ATP-BINDING DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U03853; AAA17651.1; .
CC InterPro: IPR002197; HTH_Fis.
CC InterPro: IPR002078; Sig54_interact.
CC Pfam; PF02954; HTH_8; 1.
CC Pfam; PF00158; Sigma54_activat; 1.
CC PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
CC PROSITE; PS00676; SIGMA54_INTERACT_2; FALSE_NEG.
CC PROSITE; PS00688; SIGMA54_INTERACT_3; FALSE_NEG.
CC PROSITE; PS00443; SIGMA54_INTERACT_4; 1.
CC TRANSFERASE; Glutamine amidotransferase; Complete proteome.
FT INIT_MET 0 0 BY SIMILARITY.
FT ACT_SITE 1 1 GATASE (BY SIMILARITY).
SQ SEQUENCE 314 AA; 34802 MW; A635CE2309BEB0C7 CRC64;

Query Match 1.8%; Score 7; DB 1; Length 314;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 VLEELEL 219
DB 284 VLEELEL 290
|||||||

RESULT 26
SYFA_THEME
ID SYFA_THEME STANDARD; PRT; 325 AA.
AC Q9WZ58;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phenylalanine-tRNA synthetase alpha chain (EC 6.1.1.20) (Phenylalanine-
DE tRNA ligase alpha chain) (PHERS).
GN PHERS OR TM0821.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;

```

OC Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bac

mycoplasma pneumoniae.
Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;

```
CC Mycoplasmataceae; Mycoplasma.
CC NCBI_TaxID=2104;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN-ATCC 29342 / M129;
CC MEDLINE-97105885; PubMed=8948633;
CC Himmelreich R., Hilbert H., Plagens H., Pirkle E., Li B.-C.,
CC Herrmann R.;
CC "Complete sequence analysis of the genome of the bacterium Mycoplasma
CC pneumoniae";
CC RT Nucleic Acids Res. 24:4420-4449(1996).
CC RL SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC -!- (Potential).
CC CC
CC -!- SIMILARITY: BELONGS TO THE MG185 / MG260 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000054; AAB96202.1;
CC KW Hypothetical protein; Lipoprotein; Membrane; Signal;
CC KW Complete proteome.
CC FT SIGNAL 1 22 POTENTIAL.
CC FT CHAIN 23 377 HYPOTHETICAL LIPOPROTEIN MPN281.
CC FT LIPID 23 23 N-ACYL DIGLYCERIDE (POTENTIAL).
CC SQ SEQUENCE 377 AA; 41453 MW; 8C6A7C6EE34818B5 CRC64;
CC
CC Query Match 1.8%; Score 7; DB 1; Length 377;
CC Best Local Similarity 100.0%; Pred. No. 43;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 363 TSKAAN 369
CC Db 44 TSKAAN 50
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE-90204539; PubMed=1969491;
CC Caizzi R., Bozzetti M.P., Caggese C., Ritossa F.;
CC "Homologous nuclear genes encode cytoplasmic and mitochondrial
CC glutamine synthetase in Drosophila melanogaster.";
CC J. Mol. Biol. 212:17-26(1990).
CC RT L-glutamine.
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) -> ADP + phosphate +
CC L-glutamine.
CC -!- SUBUNIT: HOMODIMER.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X52760; CAA36971.1;
CC DR PIR: S09109; AJFFIM.
CC DR FlyBase: FBgn0001142; Gsl.
CC DR InterPro: IPR001691; GLN_Synth.
CC PFam: PF00120; gln-synt_1.
CC DR PROSITE: PS00180; GLNA_1; 1.
CC DR PROSITE: PS00181; GLNA_ATP; 1.
CC KW Ligase; Multigene family; Mitochondrion; Transit peptide.
CC FT TRANSIT 1 27 MITOCHONDRION (POTENTIAL).
CC FT CHAIN 28 399 GLUTAMINE SYNTHETASE 1.
CC SQ SEQUENCE 399 AA; 44501 MW; 1A24628028268827 CRC64;
CC
CC Query Match 1.8%; Score 7; DB 1; Length 399;
CC Best Local Similarity 100.0%; Pred. No. 45;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 124 IEKLSKR 130
CC Db 310 IEKLSKR 316
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE-93043038; PubMed=1421151;
CC Brand S.N., Tan X., Widger W.R.;
CC "Cloning and sequencing of the petBD operon from the cyanobacterium
CC Synechococcus sp. PCC 7002.";
CC Plant Mol. Biol. 20:481-491(1992).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S41 (SERINE PROTEASE).
CC -!- SIMILARITY: CONTAINS 1 PDZ/DRH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X63049; CAA44776.1;
CC DR MEROPS: S41.002;
CC DR InterPro: IPR001478; PDZ.
CC DR InterPro: IPR003581; TSFC.
CC PFam: PF00595; PDZ; 1.
CC DR SMART: SM00228; PDZ; 1.
CC DR SMART: SM00245; TSFC; 1.
CC DR PROSITE: PS50106; PDZ; 1.
CC KW Hypothetical protein; Hydrolase; Serine protease.
CC FT DOMAIN 100 170 PDZ.
CC FT NON_TER 411
CC SQ SEQUENCE 411 AA; 45418 MW; B6FA904E2D7F551B CRC64;
CC
CC Query Match 1.8%; Score 7; DB 1; Length 411;
CC Best Local Similarity 100.0%; Pred. No. 46;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 130 RTVKVTR 136
CC Db 190 RTVKVTR 196
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE-93043038; PubMed=1421151;
CC Brand S.N., Tan X., Widger W.R.;
CC "Cloning and sequencing of the petBD operon from the cyanobacterium
CC Synechococcus sp. PCC 7002.";
CC Plant Mol. Biol. 20:481-491(1992).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S41 (SERINE PROTEASE).
CC -!- SIMILARITY: CONTAINS 1 PDZ/DRH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X63049; CAA44776.1;
CC DR MEROPS: S41.002;
CC DR InterPro: IPR001478; PDZ.
CC DR InterPro: IPR003581; TSFC.
CC PFam: PF00595; PDZ; 1.
CC DR SMART: SM00228; PDZ; 1.
CC DR SMART: SM00245; TSFC; 1.
CC DR PROSITE: PS50106; PDZ; 1.
CC KW Hypothetical protein; Hydrolase; Serine protease.
CC FT DOMAIN 100 170 PDZ.
CC FT NON_TER 411
CC SQ SEQUENCE 411 AA; 45418 MW; B6FA904E2D7F551B CRC64;
CC
CC Query Match 1.8%; Score 7; DB 1; Length 411;
CC Best Local Similarity 100.0%; Pred. No. 46;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 130 RTVKVTR 136
CC Db 190 RTVKVTR 196
```

```

RESULT 31
DCHS_LYCES          STANDARD;          PRT;    413 AA.
AC P54772;
AT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Histidine decarboxylase (EC 4.1.1.22) (HDC) (TOM92).
GN HDC
OS Lycopersicon esculentum (Tomato)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. AILSA CRAIG;
RX MEDLINE=94033342; PubMed=8219096;
RA Picton S., Gray J.E., Payton S., Barton S.L., Lowe A., Grierson D.;
RT "A histidine decarboxylase-like mRNA is involved in tomato fruit
RT ripening.";
RL Plant Mol. Biol. 23:627-631(1993).
CC -! CATALYTIC ACTIVITY: L-histidine = histamine + CO(2).
CC -! COFACTOR: PYRIDOXAL PHOSPHATE.
CC -! TISSUE SPECIFICITY: RIPE FRUITS; NOT DETECTED IN LEAVES AND UNRIPE
CC FRUIT.
CC -! DEVELOPMENTAL STAGE: ACCUMULATES DURING EARLY FRUIT RIPENING AND
CC THEN DECLINES.
CC -! SIMILARITY: BELONGS TO GROUP II DECARBOXYLASES (DDC, GAD, HDC AND
CC TYRDC).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X71900; CAA50719.1; -
DR InterPro; IPR002129; Pyridoxal_dec.
DR Pfam; PF00282; pyridoxal_dec; 1.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
KW Lyase; Decarboxylase; Pyridoxal phosphate.
FT BINDING 242 242 PYRIDOXAL PHOSPHATE.
SQ SEQUENCE 413 AA; 47628 MW; 5361F40B8D9A99D3 CRC64;

Query Match 1.8%; Score 7; DB 1; Length 413;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 RLEAGI 80
DB 317 RLEAGI 323
|||||

RESULT 32
IFRL_MOUSE
ID IFRL_MOUSE          STANDARD;          PRT;    449 AA.
AC P19182; P21835; P70228;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Interferon-related developmental regulator 1 (Nerve growth factor-
DE inducible protein PC4) (TPA induced sequence 7) (TIS7 protein).
GN IFRDL OR TIS7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89184573; PubMed=2467301;
RA Tirone F., Shooter E.M.;
RT "Early gene regulation by nerve growth factor in PC12 cells:
RT induction of an interferon-related gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:2088-2092(1989).
RN [2]
RP SUBCELLULAR LOCATION.
RX MEDLINE=94300650; PubMed=8028043;

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=90016295; PubMed=2797820;
RA Varum B.C., Lim R.W., Herschman H.R.;
RT "Characterization of TIS7, a gene induced in Swiss 3T3 cells by the
RT tumor promoter tetradecanoyl phorbol acetate.";
RL Oncogene 4:1263-1265(1989).
RN [2]
RP SEQUENCE OF 386-449 FROM N.A.
RX MEDLINE=82247191; PubMed=6179042;
RA Skup D., Windass J.D., Sor F., George H., Williams B.R.G.,
RA Fukuhara H., de Maeyer-Guignard J., de Maeyer E.;
RT "Molecular cloning of partial cDNA copies of two distinct mouse
RT IFN-beta mRNAs.";
RL Nucleic Acids Res. 10:3069-3084(1982).
CC -! FUNCTION: COULD PLAY A ROLE IN REGULATING GENE ACTIVITY IN THE
CC PROLIFERATIVE AND/OR DIFFERENTIATIVE PATHWAYS INDUCED BY NGF. MAY
CC BE AN AUTOCRINE FACTOR THAT ATTENUATES OR AMPLIFIES THE INITIAL
CC LIGAND-INDUCED SIGNAL.
CC -! INDUCTION: BY MITOGENS SUCH AS TPA IN 3T3 CELLS AND BY NERVE
CC GROWTH FACTOR IN PC12 PHEOCHROMOCYTOMA CELLS.
CC -! SIMILARITY: BELONGS TO THE IFRD FAMILY.
CC -! CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE INTERFERON BETA-2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17400; CAA35258.1; -
DR EMBL; V00756; CAA24133.1; ALT_SEQ.
DR PIR; A44989; A44989.
DR PIR; A05277; A05277.
DR MGD; MGI:1316717; Ifrdl.
FT DOMAIN 13 21 POLY-GLY.
FT DOMAIN 142 147 POLY-ALA.
SQ SEQUENCE 449 AA; 49943 MW; B0286C46040BA16D CRC64;

Query Match 1.8%; Score 7; DB 1; Length 449;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 EDMDSLIT 185
DB 305 EDMDSLIT 311
|||||

RESULT 33
IFRL_RAT
ID IFRL_RAT          STANDARD;          PRT;    449 AA.
AC P20695;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Interferon-related developmental regulator 1 (Nerve growth factor-
DE inducible protein PC4) (IRPR).
GN IFRD1 OR PC4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89184573; PubMed=2467301;
RA Tirone F., Shooter E.M.;
RT "Early gene regulation by nerve growth factor in PC12 cells:
RT induction of an interferon-related gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:2088-2092(1989).
RN [2]
RP SUBCELLULAR LOCATION.
RX MEDLINE=94300650; PubMed=8028043;

```

```
RA Guardavaccaro D., Montagnoli A., Ciotti M.T., Gatti A., Lotti L.,
RA di Lazzaro C., Tortisi M.R., Tirone F.;
RT "Nerve growth factor regulates the subcellular localization of the
RT nerve growth factor-inducible protein PC4 in PC12 cells.";
RL J. Neurosci. Res. 37:660-674(1994).
CC -!- FUNCTION: PROBABLY PARTICIPATES IN NEUROGENESIS. COULD PLAY A ROLE
CC IN REGULATING GENE ACTIVITY IN THE PROLIFERATIVE AND/OR
CC DIFFERENTIATIVE PATHWAYS INDUCED BY NGF.
CC -!- SUBCELLULAR LOCATION: PRESENTS A NGF-DEPENDENT PATTERN OF
CC INTRACELLULAR LOCALIZATION. WITH INCREASING AMOUNTS OF NGF AND
CC BESIDES BEING EXPRESSED IN THE CYTOPLASM, IT IS ALSO LOCALIZED IN
CC THE PLASMA MEMBRANE (INNER SIDE) AT THE ONSET OF NGF-INDUCED
CC DIFFERENTIATION, FROM WHERE IT DISAPPEARS TO REAPPEAR IN THE
CC NUCLEI OF DIFFERENTIATED CELLS.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE EMBRYONIC
CC BRAIN IN THE PERIOD RELATED TO NEUROBLAST PROLIFERATION AND
CC DIFFERENTIATION.
CC -!- DEVELOPMENTAL STAGE: ACTIVATED AT THE ONSET OF NEURONAL
CC DIFFERENTIATION.
CC -!- INDUCTION: BY NERVE GROWTH FACTOR.
CC -!- SIMILARITY: BELONGS TO THE IFRD FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J04511; AAC28946.1; -.
CC PIR; A30303; A30303.
CC Neurogenesis; Nuclear protein.
CC DOMAIN 13 21 POLY-GLY.
CC DOMAIN 142 147 POLY-ALA.
CC SEQUENCE 449 AA; 49783 MW; E1385CC622942219 CRC64;
CC -----
Query Match 1.8%; Score 7; DB 1; Length 449;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 179 EDMDSLT 185
Db 305 EDMDSLT 311
IIIIIIII
RESULT 34
COAT_FLDV STANDARD; PRT; 459 AA.
AC P22176;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Major capsid protein (MCP) (Major structural protein) (MSP).
GN MCP.
OS Fish lymphocystis disease virus (FLDV).
OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Lymphocystivirus.
OX NCBI_TaxID=10495;
RN [1]
RP PRELIMINARY SEQUENCE FROM N.A.
RC STRAIN=F;
RX MEDLINE=91220734; PubMed=2024501;
RA Schnitzler P., Handermann M., Szepe O., Darai G.;
RT "The primary structure of the thymidine kinase gene of fish
RT lymphocystis disease virus.";
RL Virology 182:835-840(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97288298; PubMed=9143276;
RA Tidona C.A., Darai G.;
RT "The complete DNA sequence of lymphocystis disease virus.";
RL Virology 230:207-216(1997).
CC -!- FUNCTION: MAJOR STRUCTURAL PROTEIN LOCATED ON THE EXTERIOR OF THE
```

```
CC CAPSID.
CC -!- SIMILARITY: BELONGS TO THE IRIDOVIRUS COAT PROTEIN FAMILY.
CC -!- CAUTION: Was originally (Ref.1) thought to be a thymidine kinase.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L63545; AAC24486.2; -.
CC PIR; A39995; K1XFFV.
CC Coat protein.
CC SEQUENCE 459 AA; 51346 MW; 4EF4C173B4C35D9B CRC64;
CC -----
Query Match 1.8%; Score 7; DB 1; Length 459;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 QKFESYF 27
Db 127 QKFESYF 133
IIIIIIII
RESULT 35
YTM1_YEAST STANDARD; PRT; 460 AA.
AC Q12024;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Microtubule-associated protein YTM1.
GN YTM1 OR YOR272W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SM201;
RA Matsumoto S., Yahara I.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S289C;
RX MEDLINE=97051594; PubMed=8896271;
RA Cheret G., Bernardi A., Sor F.J.;
RT "DNA sequence analysis of the VPH1-SNF2 region on chromosome XV of
RT Saccharomyces cerevisiae.";
RL Yeast 12:1059-1064(1996).
CC -!- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U92821; AAB51396.1; -.
CC EMBL; X89633; CAA61778.1; -.
CC EMBL; 275180; CAA99497.1; -.
CC SGD; S0005798; YTM1.
CC InterPro; IPR001680; WD40.
CC Pfam; PF00400; WD40; 6.
CC PRINTS; PRO0320; GPROTEINBRPT.
CC SMART; SM00320; WD40; 5.
CC PROSITE; PS00678; WD_REPEATS_1; 2.
CC PROSITE; PS50082; WD_REPEATS_2; 5.
CC PROSITE; PS50294; WD_REPEATS_REGION; 1.
```

```
KW Microtubules; Repeat; WD repeat.
FT REPEAT 101 140 WD 1.
FT REPEAT 142 180 WD 2.
FT REPEAT 206 244 WD 3.
FT REPEAT 285 325 WD 4.
FT REPEAT 373 413 WD 5.
FT REPEAT 424 459 WD 6.
SQ SEQUENCE 460 AA; 51358 MW; DBC26738651433C4 CRC64;

Query Match
Best Local Similarity 1.88; Score 7; DB 1; Length 460;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 170 NDKVEAV 176
DB 426 NDKVEAV 432
|||||||

RESULT 36
C13B_MYCTU STANDARD; PRT; 472 AA.
AC O53765;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative cytochrome P450 135B1 (BC 1.14.-.-).
GN CYP135B1 OR RV0568 OR MT0594 OR MT039.06.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae.
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fieischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AL021942; CAJ17439.1; -.
CC TIGR; MT0394; -.
CC DR EMBL; AE006957; AAK44817.1; -.
CC DR TIGR; MT0394; -.
CC DR InterPro; IPR001128; Cyt_P450.
CC Pfam; PF00067; p450; 2.
CC DR

KW PRINTS: PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Hypothetical protein; Oxidoreductase; Monooxygenase; Heme;
FT BINDING 388 HEME (BY SIMILARITY).
SQ SEQUENCE 472 AA; 50687 MW; B7E2898BEE80863F CRC64;

Query Match
Best Local Similarity 1.88; Score 7; DB 1; Length 472;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 LTAEEV 118
DB 326 LTAEEV 332
|||||||

RESULT 37
YU88_MYCTU STANDARD; PRT; 474 AA.
ID YU88_MYCTU
AC O53305;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 50.9 kDa protein RV3088.
GN RV3088 OR MT3173 OR MTV013.09.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae.
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fieischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE UPF0089 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AL021309; CAJ16146.1; -.
CC TIGR; MT3173; -.
CC DR EMBL; AE007134; AAK47509.1; -.
CC DR TIGR; MT3173; -.
CC DR TubercuList; RV3088; -.
CC DR InterPro; IPR004255; UPF0089.
CC Pfam; PF03007; UPF0089; 1.
CC KW Hypothetical protein; Complete proteome.
CC SEQUENCE 474 AA; 50886 MW; 36832D972BE3851A CRC64;
SQ
```



```

Query Match          1.8%; Score 7; DB 1; Length 474;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 IKGLTKL 9
DB 194 IKGLTKL 200

RESULT 38
PE24_HUMAN
ID PE24_HUMAN STANDARD; PRT; 488 AA.
AC P35408;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Prostaglandin E2 receptor, EP4 subtype (Prostanoid EP4 receptor) (PGE
DE receptor, EP4 subtype).
GN PTGER4 OR PTGER2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94216291; PubMed=8163486;
RA Bastien L., Sawyer N., Grygorczyk R., Metters K.M., Adam M.;
RT "Cloning, functional expression, and characterization of the human
RT prostaglandin E2 receptor EP2 subtype.";
RL J. Biol. Chem. 269:11873-11877(1994).
RN 2;
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94071953; PubMed=8250933;
RA An S., Yang J., Xia M., Goetzl E.J.;
RT "Cloning and expression of the EP2 subtype of human receptors for
RT prostaglandin E2.";
RL Biochem. Biophys. Res. Commun. 197:263-270(1993).
RN 3;
RP SEQUENCE FROM N.A.
RX MEDLINE=96299755; PubMed=8661119;
RA Food S.M., Marks B., Stolz M., Buffler E., Fraser N.J., Lee M.G.;
RT "The structure of the prostaglandin EP4 receptor gene and related
RT pseudogenes.";
RL Genomics 35:182-188(1996).
RN 4;
RP SEQUENCE FROM N.A.
RX MEDLINE=97015877; PubMed=8662514;
RA Mori K., Tanaka I., Kotani M., Miyaoka F., Sando T., Muro S.,
RA Sasaki Y., Nakagawa O., Ogawa Y., Usui T., Ozaki S., Ichikawa A.,
RA Narumiya S., Nakao K.;
RT "Gene expression of the human prostaglandin E receptor EP4 subtype:
RT differential regulation in monocyte and lymphoid lineage cells by
RT phorbol ester.";
RL J. Mol. Med. 74:333-336(1996).
CC -!- FUNCTION: RECEPTOR FOR PROSTAGLANDIN E2 (PGE2). THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G-S PROTEINS THAT STIMULATES
CC ADENYLATE CYCLASE. HAS A RELAXING EFFECT ON SMOOTH MUSCLE. MAY
CC PLAY AN IMPORTANT ROLE IN REGULATING RENAL HEMODYNAMICS,
CC INTESTINAL EPITHELIAL TRANSPORT, ADRENAL ALDOSTERONE SECRETION,
CC AND UTERINE FUNCTION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: HIGH IN INTESTINE AND IN PERIPHERAL BLOOD
CC MONONUCLEAR CELLS; LOW IN LUNG, KIDNEY, THYMUS, UTERUS,
CC VASCULATURE AND BRAIN. NOT FOUND IN LIVER, HEART, RETINA OE
CC SKELETAL MUSCLE.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- CAUTION: WAS ORIGINALLY DESIGNATED AS THE EP2 SUBTYPE.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L28175; AAA36434.1; -
DR EMBL; L25124; AAA36438.1; -
DR EMBL; X97873; CA66463.1; -
DR EMBL; X97874; CA66463.1; JOINED.
DR EMBL; D28472; BAA05834.1; -
DR PIR; A53572; A53572.
DR GCRDB; GCR_0836; -
DR GCRDB; GCR_0889; -
DR GCRDB; GCR_1160; -
DR MTM; 601586; -
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00428; PROTAGLNDNR.
DR PRINTS; PR00586; PRSTNOIDEP4R.
DR PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 19
FT TRANSMEM 20 43
FT DOMAIN 44 55
FT TRANSMEM 56 79
FT DOMAIN 80 96
FT TRANSMEM 97 115
FT DOMAIN 116 135
FT TRANSMEM 136 160
FT DOMAIN 161 184
FT TRANSMEM 185 211
FT DOMAIN 212 267
FT TRANSMEM 268 295
FT DOMAIN 296 312
FT TRANSMEM 313 332
FT DOMAIN 333 488
FT CARBOHYD 7
FT DISULFID 92 170
FT DOMAIN 230 234
FT CONFLICT 464 466
SQ SEQUENCE 488 AA; 53119 MW; D028478CD72C85EE CRC64;
Query Match          1.8%; Score 7; DB 1; Length 488;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 KATEKIK 329
DB 339 KATEKIK 345
-----
RESULT 39
PE24_RABIT
ID PE24_RABIT STANDARD; PRT; 488 AA.
AC Q28691;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Prostaglandin E2 receptor, EP4 subtype (Prostanoid EP4 receptor) (PGE
DE receptor, subtype EP4).
GN PTGER4.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney cortex;
RX MEDLINE=96373961; PubMed=8780252;
RA Breyer R.M., Davis L.S., Nian C., Redha R., Stillman B.,

```

RA Jacobson H.R., Breyer M.D.;
 RT "Cloning and expression of the rabbit prostaglandin EP4 receptor.";
 RL Am. J. Physiol. 270:F485-F493(1996).
 CC -!- FUNCTION: RECEPTOR FOR PROSTAGLANDIN E2 (PGE2). THE ACTIVITY OF
 CC THIS RECEPTOR IS MEDIATED BY G-S PROTEINS THAT STIMULATES
 CC ADENYLATE CYCLASE. HAS A RELAXING EFFECT ON SMOOTH MUSCLE. MAY
 CC PLAY AN IMPORTANT ROLE IN REGULATING RENAL HEMODYNAMICS.
 CC INTERSTITIAL EPITHELIAL TRANSPORT, ADRENAL ALDOSTERONE SECRETION,
 CC AND UTERINE FUNCTION.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN INTESTINE, DUODENAL
 CC EPITHELIUM, UTERUS, THYMUS AND ADRENAL CORTEX. LOWER BUT
 CC SIGNIFICANT EXPRESSION IN WHOLE ADRENAL, LUNG, SPLEEN, STOMACH
 CC AND KIDNEY. IN THIS LATTER ORGAN, THE RECEPTOR IS LOCALIZED IN THE
 CC GLOMERULI AND THE TRANSITIONAL EPITHELIUM OF THE RENAL CALYX.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L47207; AAC37330.1; -
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm1.1; 1.
 CC PRINTS; PR00428; PROTAGLNDNR.
 CC PROSITE; PR00586; PRSTNOIDEP4.
 CC PROSITE; PS00237; G-PROTEIN_RECP_FL_1; 1.
 CC PROSITE; PS00262; G-PROTEIN_RECP_FL_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein.
 KW DOMAIN 1 19 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 20 43 1 (POTENTIAL).
 FT DOMAIN 44 55 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 56 79 2 (POTENTIAL).
 FT DOMAIN 80 96 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 97 115 3 (POTENTIAL).
 FT DOMAIN 116 135 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 136 160 4 (POTENTIAL).
 FT DOMAIN 161 184 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 185 211 5 (POTENTIAL).
 FT DOMAIN 212 270 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 271 298 6 (POTENTIAL).
 FT DOMAIN 299 315 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 316 335 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 336 488 POLY-ALA.
 FT CARBOHYD 7 7 BY SIMILARITY.
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 488 AA; 53362 MW; F1217B5B0149D3C9 CRC64;
 Query Match 1.8%; Score 7; DB 1; Length 488;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 323 KAIEKIK 329
 Db 342 KAIEKIK 348
 RESULT 40
 PE24_RAT STANDARD; PRT; 488 AA.
 ID PE24_RAT
 AC P43114; O08728;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Prostaglandin E2 receptor, EP4 subtype (Prostanoid EP4 receptor) (PGE
 DE receptor, EP4 subtype).
 GN PTGER4.

OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=942411990; PubMed=8185583;
 RA Sando T., Usui T., Tanaka I., Mori K., Sasaki Y., Fukuda Y.,
 RA Namba T., Sugimoto Y., Ichikawa A., Narumiya S., Nakao K.;
 RT "Molecular cloning and expression of rat prostaglandin E receptor EP2
 RT subtype.";
 RL Biochem. Biophys. Res. Commun. 200:1329-1333(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney;
 RX MEDLINE=98141044; PubMed=9537820;
 RA Boie Y., Stocco R., Sawyer N., Slipetz D.M., Ungrin M.D.,
 RA Neuschaefer-Rube F., Puschel G.P., Metters K.M., Abramovitz M.;
 RT "Molecular cloning and characterization of the four rat prostaglandin
 RT E2 prostanoid receptor subtypes.";
 RL Eur. J. Pharmacol. 340:227-241(1997).
 CC -!- FUNCTION: RECEPTOR FOR PROSTAGLANDIN E2 (PGE2). THE ACTIVITY OF
 CC THIS RECEPTOR IS MEDIATED BY G-S PROTEINS THAT STIMULATES
 CC ADENYLATE CYCLASE. HAS A RELAXING EFFECT ON SMOOTH MUSCLE. MAY
 CC PLAY AN IMPORTANT ROLE IN REGULATING RENAL HEMODYNAMICS,
 CC INTERSTITIAL EPITHELIAL TRANSPORT, ADRENAL ALDOSTERONE SECRETION,
 CC AND UTERINE FUNCTION.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -!- CAUTION: WAS ORIGINALLY DESIGNATED AS THE EP2 SUBTYPE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D28860; BAA06011.1; -
 CC EMBL; U94709; AAB53326.1; -
 CC GCRDB; GCR_1275; -
 CC GCRDB; GCR_1496; -
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm1.1; 1.
 CC PRINTS; PR00428; PROTAGLNDNR.
 CC PRINTS; PR00586; PRSTNOIDEP4.
 CC PROSITE; PS00237; G-PROTEIN_RECP_FL_1; 1.
 CC PROSITE; PS00262; G-PROTEIN_RECP_FL_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein.
 KW DOMAIN 1 19 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 20 43 1 (POTENTIAL).
 FT DOMAIN 44 55 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 56 79 2 (POTENTIAL).
 FT DOMAIN 80 96 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 97 115 3 (POTENTIAL).
 FT DOMAIN 116 135 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 136 160 4 (POTENTIAL).
 FT DOMAIN 161 184 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 185 211 5 (POTENTIAL).
 FT DOMAIN 212 270 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 271 298 6 (POTENTIAL).
 FT DOMAIN 299 315 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 316 335 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 336 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 7 7 BY SIMILARITY.
 FT CARBOHYD 92 92 L -> H (IN REF. 2).
 FT CARBOHYD 396 396 H -> O (IN REF. 2).
 SO SEQUENCE 488 AA; 53365 MW; C2EE713018723FF4 CRC64;

```
Query Match 1.8%, Score 7; DB 1; Length 488;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 KAIEKIK 329
    |||||
Db 342 KAIEKIK 348

RESULT 41
ID A2AP_HUMAN STANDARD; PRT; 491 AA.
AC P08697;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Alpha-2-antiplasmin precursor (Alpha-2-plasmin inhibitor) (Alpha-2-PI)
DE (Alpha-2-AP).
GN SERPINF2 OR PLI OR AAP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88139254; PubMed=2830248;
RA Tone M., Kikuno R., Kume-Twaki A., Hashimoto-Gotoh T.;
RT "Structure of human alpha 2-plasmin inhibitor deduced from the cDNA
sequence.";
RL J. Biochem. 102:1033-1041(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88320531; PubMed=3166140;
RA Hirosewa S., Nakamura Y., Miura O., Sumi Y., Aoki N.;
RT "Organization of the human alpha 2-plasmin inhibitor gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:6836-6840(1988).
RN [3]
RP ERRATUM.
RA Hirosewa S., Nakamura Y., Miura O., Sumi Y., Aoki N.;
RL Proc. Natl. Acad. Sci. U.S.A. 86:1612-1613(1989).
RN [4]
RP SEQUENCE OF 4-491 FROM N.A.
RX MEDLINE=87109313; PubMed=2433286;
RA Holmes W.E., Nelles L., Lijnen H.R., Collen D.;
RT "Primary structure of human alpha 2-antiplasmin, a serine protease
inhibitor (serpin).";
RL J. Biol. Chem. 262:1659-1664(1987).
RN [5]
RP SEQUENCE OF 218-491 FROM N.A.
RX MEDLINE=87137400; PubMed=3818581;
RA Sumi Y., Nakamura Y., Aoki N., Sakai M., Muramatsu M.;
RT "Structure of the carboxyl-terminal half of human alpha 2-plasmin
inhibitor deduced from that of cDNA.";
RL J. Biochem. 100:1399-1402(1986).
RN [6]
RP SEQUENCE OF 40-491.
RX MEDLINE=87275946; PubMed=2440681;
RA Lijnen H.R., Holmes W.E., van Hoef B., Wiman B., Rodriguez H.,
Collen D.;
RT "Amino acid sequence of human alpha 2-antiplasmin.";
RL Eur. J. Biochem. 166:565-574(1987).
RN [7]
RP SEQUENCE OF 40-43.
RX MEDLINE=78023887; PubMed=21075;
RA Wiman B., Collen D.;
RT "Purification and characterization of human antiplasmin, the
fast-acting plasmin inhibitor in plasma.";
RL Eur. J. Biochem. 78:19-26(1977).
RN [8]
RP SEQUENCE OF 28-52.
RX TISSUE=Plasma;
RL MEDLINE=93050153; PubMed=1385210;
RA Christensen S., Sottrup-Jensen L.;
```

```
RT "Bovine alpha 2-antiplasmin. N-terminal and reactive site sequence.";
RL FEBS Lett. 312:100-104(1992).
RN [9]
RP ACTIVE SITES.
RX MEDLINE=88290696; PubMed=2456616;
RA Potempa J., Shieh B.-H., Travis J.;
RT "Alpha-2-antiplasmin: a serpin with two separate but overlapping
reactive sites.";
RL Science 241:699-700(1988).
RN [10]
RP SEQUENCE OF 481-491, AND SULFATION.
RX MEDLINE=87137577; PubMed=2434496;
RA Hordt G., Fok K.F., Toren P.C., Strauss A.W.;
RT "Sulfation of a tyrosine residue in the plasmin-binding domain of
alpha 2-antiplasmin.";
RL J. Biol. Chem. 262:3082-3085(1987).
CC -!- FUNCTION: THE MAJOR TARGETS OF THIS INHIBITOR ARE PLASMIN AND
TRYPSIN, BUT IT ALSO INACTIVATES CHYMOTRYPSIN.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL: D00116; BAA00070.1; -
CC DR EMBL: D00174; BAA00124.1; -
CC DR EMBL: M20786; AAA51554.1; -
CC DR EMBL: M20782; AAA51554.1; JOINED.
CC DR EMBL: M20783; AAA51554.1; JOINED.
CC DR EMBL: M20784; AAA51554.1; JOINED.
CC DR EMBL: M20785; AAA51554.1; JOINED.
CC DR EMBL: J02654; AAA35543.1; -
CC DR PIR: A24708; A24708.
CC DR PIR: A26684; A26684.
CC DR PIR: A31402; A31402.
CC DR PIR: A32163; A32163.
CC DR PIR: A41504; A41504.
CC DR PIR: S00068; S00068.
CC DR HSSP: P01008; LANT.
CC DR SWISS-2DPAGE; P08697; HUMAN.
CC DR MIM; 262850; -
CC DR InterPro; IPR000215; Serpin.
CC DR Pfam; PF00079; serpin; 1.
CC DR SMART; SM00093; SERPIN; 1.
CC DR PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor; Glycoprotein; Plasma; Signal;
KW Sulfation.
FT SIGNAL 1 27
FT PROPEP 28 39
FT CHAIN 40 491
FT ACT_SITE 403 404
FT ACT_SITE 404 405
FT MOD_RES 484 484
FT CARBOHYD 126 126
FT CARBOHYD 295 295
FT CARBOHYD 309 309
FT CARBOHYD 316 316
FT CONFLICT 33 33
FT CONFLICT 49 49
FT CONFLICT 105 105
FT CONFLICT 289 289
FT CONFLICT 408 408
FT CONFLICT 455 455
FT CONFLICT 491 AA; 54565 MW; 385A1C90E91A63CB CRC64;
SQ SEQUENCE 491 AA; 54565 MW; 385A1C90E91A63CB CRC64;

Query Match 1.8%; Score 7; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 243 GQQTALK 249
DB 58 GQQTALK 64

RESULT 42
GTRL_MOUSE
ID GTRL_MOUSE STANDARD; PRT; 492 AA.
AC P17809; Q61608;
DT 01-AUG-1990 (Rel. 15, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 1
DE SLC2A1 OR GLUT1 OR GLUT-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=69240694; PubMed=2654938;
RA Kaestner K.H., Christy R.J., McLenithan J.C., Braiterman L.T.,
RA Cornelius P., Pekala P.H., Lane M.D.;
RT "Sequence, tissue distribution, and differential expression of mRNA
RT for a putative insulin-responsive glucose transporter in mouse 3T3-L1
RT adipocytes.";
RL Arch. Biochem. Biophys. 279:261-274(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90274408; PubMed=2190533;
RA Reed B.C., Shade D., Alperovich F., Vang M.;
RT "3T3-L1 adipocyte glucose transporter (HepG2 class): sequence and
RT regulation of protein and mRNA expression by insulin,
RT differentiation, and glucose starvation.";
RL Arch. Biochem. Biophys. 279:261-274(1990).
RN [3]
RP SEQUENCE OF 151-237 FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Embryo;
RX MEDLINE=93170163; PubMed=1289053;
RA Smith D.E., Gridley T.;
RT "Differential screening of a PCR-generated mouse embryo cDNA library:
RT glucose transporters are differentially expressed in early
RT postimplantation mouse embryos.";
RL Development 116:555-561(1992).
RN [4]
RP SEQUENCE OF 357-463 FROM N.A.
RX MEDLINE=92111400; PubMed=1765007;
RA Hogan A., Heyner S., Charton M.J., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A., Thorens B., Schultz G.A.;
RT "Glucose transporter gene expression in early mouse embryos.";
RL Development 113:363-372(1991).
CC -!- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM MAY BE
CC RESPONSIBLE FOR CONSTITUTIVE OR BASAL GLUCOSE UPTAKE. HAS A VERY
CC BROAD SUBSTRATE SPECIFICITY; CAN TRANSPORT A WIDE RANGE OF ALDOSES
CC INCLUDING BOTH PENTOSE AND HEXOSE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localizes
CC primarily at the cell surface (By similarity).
CC -!- DEVELOPMENTAL STAGE: LEVELS DECLINE 3-FOLD BETWEEN DAYS 7.5 AND
CC 12.5 OF GESTATION. AT 7.5 DPC, EXPRESSED MORE STRONGLY IN
CC EXTRAEMBRYONIC TISSUES THAN IN THE EMBRYO PROPER. EXPRESSED IN
CC AMNION, CHORION, AND ECTOPLACENTAL CONE. IN THE YOLK SAC,
CC EXPRESSED MORE STRONGLY IN THE MESODERM LAYER THAN THE ECTODERM.
CC EXPRESSION FAIRLY WIDESPREAD IN THE EMBRYO AT 8.5 DPC, BUT BY 10.5
CC DPC, EXPRESSION IS DOWN-REGULATED AND OBSERVED IN THE EYE AND THE
CC SPINAL CORD.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC TRANSPORTER SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; M23384; AAA37752.1; -
DR EMBL; M22998; AAA37707.1; -
DR EMBL; X69697; CAA49367.1; -
DR EMBL; S77924; AAB20846.2; -
DR PIR; A30310; A30310.
DR PIR; S09705; S09705.
DR HSSP; P00844; IAg1.
DR MGI; MGI:95755; Sic2al.
DR InterPro; IPR003663; Sugar_transporter.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Transport; Glycoprotein;
KW Multigene family.
FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 13 33 1 (POTENTIAL).
FT DOMAIN 34 66 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 67 87 2 (POTENTIAL).
FT DOMAIN 88 95 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 96 116 3 (POTENTIAL).
FT DOMAIN 117 126 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 127 147 4 (POTENTIAL).
FT DOMAIN 148 155 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 156 176 5 (POTENTIAL).
FT DOMAIN 177 185 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 186 206 6 (POTENTIAL).
FT DOMAIN 207 271 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 272 292 7 (POTENTIAL).
FT DOMAIN 293 307 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 308 328 8 (POTENTIAL).
FT DOMAIN 329 337 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 338 358 9 (POTENTIAL).
FT DOMAIN 359 371 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 372 392 10 (POTENTIAL).
FT DOMAIN 393 401 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 402 422 11 (POTENTIAL).
FT DOMAIN 423 429 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 430 450 12 (POTENTIAL).
FT DOMAIN 451 492 CYTOPLASMIC (POTENTIAL).
FT SITE 279 281 DEFINES SUBSTRATE SPECIFICITY (BY SIMILARITY).
FT CARBOHYD 45 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 52 I -> Y (IN REF. 2).
FT CONFLICT 193 IFI -> VFV (IN REF. 1).
FT CONFLICT 357 LLER -> MQEQ (IN REF. 4).
FT CONFLICT 403 A -> R (IN REF. 1).
SQ SEQUENCE 492 AA; 53934 MW; 42106466B26F3735 CRC64;

Query Match 1.8%; Score 7; DB 1; Length 492;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 MDPSSKK 202
DB 1 MDPSSKK 7

RESULT 43
YFCC_ECOLI STANDARD; PRT; 506 AA.
ID YFCC_ECOLI
AC P39263; P76493; P77678;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yfcc.
GN YFCC OR B2298.

```

OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubramam S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT - K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features";
RL DNA Res. 4:91-113(1997).
RN [3]
RP SEQUENCE OF 1-90 FROM N.A.
RC STRAIN=K12 / KHI31;
RX MEDLINE=95189796; PubMed=7883769;
RA Kakuda H., Hosono K., Shiroishi K., Ichihara S.;
RT "Identification and characterization of the ackA (acetate kinase
RT A)-pta (phosphotransacetylase) operon and complementation analysis of
RT acetate utilization by an ackA-pta deletion mutant of Escherichia
RT coli";
RL J. Biochem. 116:916-922(1994).
RN [4]
RP IDENTIFICATION.
RA Rudd K.E.;
RL Unpublished observations (NOV-1994).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: TO H.INFLUENZAE HI0594.
CC -!- SIMILARITY: TO B.SUBTILIS YCGA.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000319; AAC75358.1; ALT_INIT.
CC EMBL; D90861; BAA16137.1; -.
CC EMBL; D90862; BAA16146.1; -.
CC EMBL; D17376; -; NOT_ANNOTATED_CDS.
CC Ecogene: Esi2607; Yfccc.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 107 127 POTENTIAL.
FT TRANSMEM 148 168 POTENTIAL.
FT TRANSMEM 171 191 POTENTIAL.
FT TRANSMEM 208 228 POTENTIAL.
FT TRANSMEM 231 251 POTENTIAL.
FT TRANSMEM 287 307 POTENTIAL.
FT TRANSMEM 310 330 POTENTIAL.
FT TRANSMEM 352 372 POTENTIAL.
FT TRANSMEM 398 418 POTENTIAL.
FT TRANSMEM 419 439 POTENTIAL.
FT TRANSMEM 442 462 POTENTIAL.
FT TRANSMEM 485 505 POTENTIAL.

FT CONFLICT 1 34 MSATKSPRRRWAMPDTLVIFVAILTSLATW ->
FT CONFLICT 86 90 MLFNQPCHL (IN REF. 3).
FT GDERP -> AMNAR (IN REF. 3).
SQ SEQUENCE 506 AA; 54796 MW; FA9D2244C1971D5 CRC64;
Query Match 1.8%; Score 7; DB 1; Length 506;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 54 LTNEAGE 60
IIIIIIII
DB 68 LTNEAGE 74
RESULT 44
RECN_CAMJE STANDARD; PRT; 507 AA.
ID RECN_CAMJE
AC Q9PHM5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA repair protein recN (Recombination protein N).
GN RECN OR CJ0642.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Tagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences";
RL Nature 403:665-668(2000).
CC -!- FUNCTION: MAY BE INVOLVED IN RECOMBINATIONAL REPAIR OF DAMAGED
CC DNA (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE RECN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL139075; CAB5278.1; -.
CC InterPro: IPR003439; ABC_transportr.
KW DNA repair; ATP-binding; Complete proteome.
FT NP_BIND 29 36 ATP (POTENTIAL).
SQ SEQUENCE 507 AA; 58522 MW; 7A99ABA35C268F5D CRC64;
Query Match 1.8%; Score 7; DB 1; Length 507;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 211 AKVLEEL 217
IIIIIIII
DB 432 AKVLEEL 438
RESULT 45
PE24_MOUSE
ID PE24_MOUSE STANDARD; PRT; 513 AA.
AC P32240;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)

Prostaglandin E2 receptor, EP4 subtype (Prostanoid EP4 receptor) (PGE2 receptor, EP4 subtype).
PTGER4 OR PTGEREP4.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
STRAIN=BDF1;
MEDLINE=93216741; PubMed=83951118;
Honda A., Sugimoto Y., Namba T., Watabe A., Irie A., Negishi M., Narumiya S., Ichikawa A.;
"Cloning and expression of a cDNA for mouse prostaglandin E receptor EP2 subtype";
J. Biol. Chem. 268:7759-7762(1993).
[2]
SEQUENCE OF 1-317 FROM N.A.
MEDLINE=97094651; PubMed=8939885;
Arakawa T., Laneuville O., Miller C.A., Lakkides K.M.,
Wingard B.A., Dewitt D.L., Smith W.L.;
"Prostanoid receptors of murine NIH 3T3 and RAW 264.7 cells.
Structure and expression of the murine prostaglandin EP4 receptor gene";
J. Biol. Chem. 271:29569-29575(1996).
-!- FUNCTION: RECEPTOR FOR PROSTAGLANDIN E2 (PGE2). THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G-S PROTEINS THAT STIMULATES ADENYLATE CYCLASE. HAS A RELAXING EFFECT ON SMOOTH MUSCLE. MAY PLAY AN IMPORTANT ROLE IN REGULATING RENAL HEMODYNAMICS, INTESTINAL EPITHELIAL TRANSPORT, ADRENAL ALDOSTERONE SECRETION, AND UTERINE FUNCTION.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- TISSUE SPECIFICITY: ABUNDANT EXPRESSION IN ILEUM, THYMUS AND MASTOCYTOMA P-815 CELLS. ALSO OBSERVED IN LUNG, SPLEEN, HEART AND UTERUS.
-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-!- CAUTION: WAS ORIGINALLY DESIGNATED AS THE EP2 SUBTYPE.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcements> or send an email to license@isb-sib.ch).

EMBL: D13458; BAA02714.1; -;
DR EMBL: U70365; AAC52897.1; -;
DR PIR: A46638; A46638.
DR GCRDB: GCR_0718; -;
DR MGDB: MGI:104311; Ptger4.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1.1.
DR PRINTS: PR00428; PROSTA_GLNDR.
DR PRINTS: PR00586; PRSTNOIDEP4R.
DR PROSITE: PS00237; G-PROTEIN_RECEP_FL1_1.
DR PROSITE: PS0262; G-PROTEIN_RECEP_FL2_1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein.
KW DOMAIN 1 44
FT TRANSMEM 45 68
FT DOMAIN 69 80
FT TRANSMEM 81 104
FT DOMAIN 105 121
FT TRANSMEM 122 140
FT DOMAIN 141 160
FT TRANSMEM 161 185
FT DOMAIN 186 209
FT TRANSMEM 210 236
FT DOMAIN 237 295
FT TRANSMEM 296 323
FT DOMAIN 324 340
FT TRANSMEM 341 360
FT DOMAIN 361 513

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 13:53:51 ; Search time 69 seconds

(without alignments)
950.219 Million cell updates/sec

Title: US-09-805-311-2

Perfect score: 379

Sequence: 1 MGKGLTKLLADNAPKAMKE.....SDKTSKAAANKTKAGGKKK 379

Scoring table: OLIGO

Gapop 60.0 , Capext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_19.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	16.4	380	10 Q9SXQ6	Q9SXQ6 oryza sativ
2	28	7.4	362	10 O65251	O65251 arabidopsis
3	15	4.0	343	17 Q9VOP9	Q9VOP9 pyrococcus
4	14	3.7	382	5 Q9N312	Q9N312 caenorhabdi
5	14	3.7	385	5 O96830	O96830 drosophila
6	13	3.4	378	11 Q99M86	Q99M86 mus musculu
7	13	3.4	380	11 Q91250	Q91250 mus musculu
8	13	3.4	381	13 Q90YB0	Q90YB0 gallus gall
9	13	3.4	382	13 P70040	P70040 xenopus lae
10	13	3.4	382	13 O57351	O57351 xenopus lae
11	13	3.4	382	13 P70054	P70054 xenopus lae
12	13	3.4	401	17 Q91FY5	Q91FY5 aeropyrum p
13	12	3.2	340	1 O93634	O93634 pyrococcus
14	11	2.9	380	11 Q9JHW7	Q9JHW7 rattus norv
15	10	2.6	650	5 Q9GZ01	Q9GZ01 plasmodium
16	10	2.6	672	5 Q9U0K1	Q9U0K1 plasmodium

17	9	2.4	336	17 Q9HJD4	Q9HJD4 thermoplas
18	9	2.4	378	3 Q94305	Q94305 schizosacch
19	8	2.1	155	17 Q97727	Q97727 thermoplas
20	8	2.1	278	2 Q93MR6	Q93MR6 bordetella
21	8	2.1	278	2 Q93MR5	Q93MR5 bordetella
22	8	2.1	282	2 Q932T2	Q932T2 bordetella
23	8	2.1	302	17 Q980U8	Q980U8 sulfolobus
24	8	2.1	304	17 Q976H6	Q976H6 sulfolobus
25	8	2.1	343	17 O50123	O50123 pyrococcus
26	8	2.1	354	3 Q96UV5	Q96UV5 hebeloma cy
27	8	2.1	373	11 Q9EQP8	Q9EQP8 cricetus
28	8	2.1	399	5 Q96770	Q96770 drosophila
29	8	2.1	399	5 Q9VFK0	Q9VFK0 drosophila
30	8	2.1	573	10 Q9AMB2	Q9AMB2 glycine max
31	8	2.1	770	10 Q9LDM8	Q9LDM8 glycine max
32	8	2.1	802	16 Q98CQ1	Q98CQ1 rhizobium l
33	8	2.1	862	5 Q27446	Q27446 artemia sal
34	7	1.8	59	16 Q98FF9	Q98FF9 rhizobium l
35	7	1.8	63	16 Q9JYG3	Q9JYG3 neisseria m
36	7	1.8	65	2 Q32527	Q32527 escherichia
37	7	1.8	65	9 Q92WZ6	Q92WZ6 mycobacteri
38	7	1.8	67	16 Q9K880	Q9K880 bacillus ha
39	7	1.8	68	15 Q9QJD7	Q9QJD7 human immun
40	7	1.8	78	9 Q37945	Q37945 bacterioph
41	7	1.8	96	4 Q9P192	Q9P192 homo sapien
42	7	1.8	97	16 Q91449	Q91449 pseudomonas
43	7	1.8	103	17 Q9YCG8	Q9YCG8 aeropyrum p
44	7	1.8	118	10 Q91I81	Q91I81 arabidopsis
45	7	1.8	128	10 Q9SU26	Q9SU26 arabidopsis

ALIGNMENTS

RESULT 1

Q9SXQ6 ID Q9SXQ6 PRELIMINARY; PRT; 380 AA.
AC Q9SXQ6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE FEN-1.
GN OSPEN-1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20256470; PubMed=10798612;
RA Kimura S., Ueda T., Hatanaka M., Takenouchi M., Hashimoto J.,
RA Sakaguchi K.;
RT "Plant homologue of flap endonuclease-1: molecular cloning,
RT characterization, and evidence of expression in meristematic
RT tissues."
RL Plant Mol. Biol. 42:415-427(2000).
DR EMBL: AB021666; BAA36171.1;
DR HSP: Q58839; IAT76.
DR InterPro: IPR002421; 5_3_exonuclease.
DR InterPro: IPR000513; Exo_N_I.
DR InterPro: IPR003584; HHH_2.
DR InterPro: IPR001532; XPG_I.
DR Pfam: PF01367; 5_3_exonuclease; 1.
DR Pfam: PF00867; XPG_I; 1.
DR Pfam: PF00752; XPG_N; 1.
DR PRINTS: PR00853; XPGRADSUPER.
DR SMART: SM00475; 53EXOC; 1.
DR SMART: SM00279; Hnh2; 1.
DR SMART: SM00484; XPGI; 1.
DR SMART: SM00485; XPGN; 1.
DR PROSITE: PS00841; XPG_1; 1.
SQ SEQUENCE 380 AA; 42792 MW; E0148AAFA95A7364 CRC64;

```
Query Match          16.4%; Score 62; DB 10; Length 380;
Best Local Similarity 100.0%; Pred. No. 3.8e-55;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 IAVDASMSIYQPLIVVGTGNETLTNEAGEVTSHLQGMFNRTIRLLPAGIKPVYVFDGKP 90
DB 31 IAVDASMSIYQPLIVVGTGNETLTNEAGEVTSHLQGMFNRTIRLLPAGIKPVYVFDGKP 90
QY 91 PD 92
DB 91 PD 92

RESULT 2
O65251 PRELIMINARY; PRT; 362 AA.
ID O65251
AC O65251
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 18, Last annotation update)
DE F21E10.3 PROTEIN.
GN F21E10.3
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Davidson S., Rohlfing T., David M., O'Brian D.;
RT "The sequence of A. thaliana F21E10.";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Washu;
RT "The A. thaliana Genome Sequencing Project.";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Waterston R.;
RT "Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Wilson R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF058914; AAC13596.1;
DR HSSP; Q58839; 1A76.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR000513; EXO_N_I.
DR InterPro; IPR003584; HHH_2.
DR InterPro; IPR001532; XPG_I.
DR Pfam; PF01367; 5_3_exonuclease; 1.
DR Pfam; PF00867; XPG_I; 1.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS; PR00853; XPGRADSUPER.
DR SMART; SM00279; HhH2; 1.
DR SMART; SM00484; XPGI; 1.
DR SMART; SM00485; XPGN; 1.
DR PROSITE; PS00841; XPG_1; 1.
SQ SEQUENCE 362 AA; 41205 MW; A4DDI706C3AD2D9D CRC64;

Query Match          7.48; Score 28; DB 10; Length 362;
Best Local Similarity 100.0%; Pred. No. 4.3e-20;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 HLQGMFNRTIRLLPAGIKPVYVFDGKPP 91
DB 97 HLQGMFNRTIRLLPAGIKPVYVFDGKPP 124

Query Match          4.0%; Score 15; DB 17; Length 343;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 EAGIKPVYVFDGKPP 91
DB 70 EAGIKPVYVFDGKPP 84

RESULT 4
Q9N3T2 PRELIMINARY; PRT; 382 AA.
ID Q9N3T2
AC Q9N3T2
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE HYPOTHETICAL 42.5 KDA PROTEIN.
GN Y47G6A.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Minx P., Graves T., Hawrysko C.;
RT "The sequence of C. elegans cosmid Y47G6A.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
```


RC STRAIN-BRISTOL N2;
 RA Waterston R.;
 RT "direct Submission.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC024791; AAF60653.1; -;
 DR HSSP; Q58839; 1A76.
 DR InterPro; IPR002421; 5_3_exonuclease.
 DR InterPro; IPR000513; EXO_N_I.
 DR InterPro; IPR003584; HHL_2.
 DR InterPro; IPR001532; XPG_I.
 DR Pfam; PF01367; 5_3_exonuclease; 1.
 DR Pfam; PF00867; XPG_I; 1.
 DR Pfam; PF00752; XPG_N; 1.
 DR PRINTS; PR00853; XPGRADSUPER.
 DR SMART; SM00279; HhH2; 1.
 DR SMART; SM00484; XPGI; 1.
 DR SMART; SM00485; XPGN; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 382 AA; 42549 MW; FDFE0BA0707321EB CRC64;

Query Match 3.7%; Score 14; DB 5; Length 382;
 Best Local Similarity 100.0%; Pred. No. 1.2e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 KPVYVFDGKPPDMK 94
 Db |||||
 Db 80 KPVYVFDGKPPDMK 93

RESULT 5
 O96830
 ID O96830 PRELIMINARY; PRT; 385 AA.
 AC O96830;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE EG:EG0003.3 PROTEIN.
 GN EG:EG0003.3 OR CG8648.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananathides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
 RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes N., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karp G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith I.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Murphy L., Harris D., Barrell B.;
 RT "Sequencing the distal X chromosome of Drosophila melanogaster.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Benos P.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003805; AAF57944.1; -;
 DR EMBL; AL031863; CAA21320.1; -;
 DR HSSP; Q58839; 1A76.
 DR FlyBase; FBgn025832; Fen1.
 DR InterPro; IPR002421; 5_3_exonuclease.
 DR InterPro; IPR000513; EXO_N_I.
 DR InterPro; IPR003583; HHL_1.
 DR InterPro; IPR003584; HHL_2.
 DR InterPro; IPR001532; XPG_I.
 DR Pfam; PF01367; 5_3_exonuclease; 1.
 DR Pfam; PF00867; XPG_I; 1.
 DR Pfam; PF00752; XPG_N; 1.
 DR PRINTS; PR00853; XPGRADSUPER.
 DR SMART; SM00279; HhH1; 1.
 DR SMART; SM00279; HhH2; 1.
 DR SMART; SM00484; XPGI; 1.
 DR SMART; SM00485; XPGN; 1.
 DR PROSITE; PS00841; XPG_1; 1.
 DR PROSITE; PS00842; XPG_2; 1.
 SQ SEQUENCE 385 AA; 42948 MW; D0A0831C2BDA9240 CRC64;

Query Match 3.7%; Score 14; DB 5; Length 385;
 Best Local Similarity 100.0%; Pred. No. 1.3e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVYVFDGKPPD 92
 Db |||||
 Db 78 GIKPVYVFDGKPPD 91

RESULT 6
 Q99M86
 ID Q99M86 PRELIMINARY; PRT; 378 AA.
 AC Q99M86;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE FLAP ENDONUCLEASE-1.
 GN FEN1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10909;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129SV/J;
 RA Karamjawa Z.E., Shi X., Hsieh C.-L., Lieber M.R.;
 RT "The Mammalian Fen1 Locus: Structure and Conserved Sequence Features.";
 RL Microb. Comp. Genomics 0:0-0(2001).
 DR EMBL; AY014962; AAK01853.1; -;
 DR HSSP; Q58839; 1A76.

DR InterPro; IPR002421; 5_3-exonuclease.
 DR InterPro; IPR000513; Exo_N_I.
 DR InterPro; IPR003584; HHH_2.
 DR InterPro; IPR001532; XPG_I.
 DR Pfam; PF01367; 5_3-exonuclease; 1.
 DR Pfam; PF00867; XPG_I; 1.
 DR Pfam; PF00752; XPG_N; 1.
 DR PRINTS; PR00853; XPGRADSUPER.
 DR SMART; SM00279; HhH2; 1.
 DR SMART; SM00484; XPGI; 1.
 DR SMART; SM00485; XPGN; 1.
 DR PROSITE; PS00841; XPG_1; 1.
 DR PROSITE; PS00842; XPG_2; 1.
 KW Endonuclease.
 SQ SEQUENCE 378 AA; 42300 MW; 827946BA8CAC9F39 CRC64;

Query Match 3.4%; Score 13; DB 11; Length 378;
 Best Local Similarity 100.0%; Pred. No. 0.00013;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVYVFDGKPP 91
 |||||

Db 76 GIKPVYVFDGKPP 88
 |||||

RESULT 7

Q91Z50 ID Q91Z50 PRELIMINARY; PRT; 380 AA.

AC Q91Z50;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE SIMILAR TO FLAG STRUCTURE-SPECIFIC ENDONUCLEASE.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;

RN [1]
 RP SEQUENCE FROM N.A.

RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC010203; AAH10203.1; -

KW Endonuclease

SQ SEQUENCE 380 AA; 42623 MW; 1BE903288846520D CRC64;

Query Match 3.4%; Score 13; DB 11; Length 380;
 Best Local Similarity 100.0%; Pred. No. 0.00013;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVYVFDGKPP 91
 |||||

Db 78 GIKPVYVFDGKPP 90
 |||||

RESULT 8

Q90YB0 ID Q90YB0 PRELIMINARY; PRT; 381 AA.

AC Q90YB0;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE FEN-1 NUCLEASE.

GN FEN1.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.

OC NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RA Matsuzaki Y., Adachi N., Koyama H.;

RT "The FEN-1 nuclease is not essential for DNA replication, but is
 required for maintenance of genomic integrity and base excision repair

RT in vertebrate cells.";
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB058602; BAB68507.1; -

SQ SEQUENCE 381 AA; 43054 MW; 71432ECBCBAE67AA CRC64;

Query Match 3.4%; Score 13; DB 13; Length 381;
 Best Local Similarity 100.0%; Pred. No. 0.00013;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVYVFDGKPP 91
 |||||

Db 78 GIKPVYVFDGKPP 90
 |||||

RESULT 9

P70040 ID P70040 PRELIMINARY; PRT; 382 AA.

AC P70040;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)

DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)

DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)

DE 5' NUCLEASE XFEN1A.

GN FEN1.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

OC Xenopodinae; Xenopus.

OC NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RA Bibikova M., Chi E., Wu B., Kim K., Carroll D.;

RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Li J.-L., Cox L.S.;

RT "Cloning and investigation of Xenopus Fen1: developmental expression
 and function in DNA replication.";

RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; U64563; AAB06176.1; -

DR EMBL; AF065397; AAD02814.1; -

DR HSSP; O58839; 1A76.

DR InterPro; IPR002421; 5_3-exonuclease.

DR InterPro; IPR000513; Exo_N_I.

DR InterPro; IPR003584; HHH_2.

DR InterPro; IPR001532; XPG_I.

DR Pfam; PF01367; 5_3-exonuclease; 1.

DR Pfam; PF00867; XPG_I; 1.

DR Pfam; PF00752; XPG_N; 1.

DR PRINTS; PR00853; XPGRADSUPER.

DR SMART; SM00279; HhH2; 1.

DR SMART; SM00484; XPGI; 1.

DR SMART; SM00485; XPGN; 1.

KW Endonuclease.

SQ SEQUENCE 382 AA; 42668 MW; 9B1DB0EDAD158D57 CRC64;

Query Match 3.4%; Score 13; DB 13; Length 382;
 Best Local Similarity 100.0%; Pred. No. 0.00013;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVYVFDGKPP 91
 |||||

Db 78 GIKPVYVFDGKPP 90
 |||||

RESULT 10

O57351 ID O57351 PRELIMINARY; PRT; 382 AA.

AC O57351;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)

DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE FLAP ENDONUCLEASE 1.

GN FEN1.

```

OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98204872; PubMed=9535864;
RA Kim K., Blade S., Matsumoto Y.;
RT "Involvement of flap endonuclease 1 in base excision DNA repair.";
RL J. Biol. Chem. 273:8842-8848(1998).
DR EMBL; AF036327; AB888707.1; -.
DR HSSP; Q58839; 1A76.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR000513; EXO_N_I.
DR InterPro; IPR003584; HHH_2.
DR InterPro; IPR001532; XPG_I.
DR Pfam; PF01367; 5_3_exonuclease; 1.
DR Pfam; PF00867; XPG_I; 1.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS; PR00853; XPGRADSUPER.
DR SMART; SM00279; HHH2; 1.
DR SMART; SM00484; XPGI; 1.
DR SMART; SM00485; XPGN; 1.
KW Endonuclease.
SQ SEQUENCE 382 AA; 42682 MW; 3A911F83DC5B5A59 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVVFDGKPP 91
DB 78 GIKPVVFDGKPP 90
|||||

RESULT 11
P70054
ID P70054 PRELIMINARY; PRT; 382 AA.
AC P70054;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE XFN1B.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=HEAD;
RX MEDLINE=99069415; PubMed=9852084;
RA Bibikova M., Wu B., Chi E., Kim K.H., Trautman J.K., Carroll D.;
RT "Characterization of FEN-1 from Xenopus laevis. cDNA cloning and role
in DNA metabolism.";
RL J. Biol. Chem. 273:34222-34229(1998).
DR EMBL; U68141; AAB08478.1; -.
DR HSSP; Q58839; 1A76.
DR InterPro; IPR002421; 5_3_exonuclease.
DR InterPro; IPR000513; EXO_N_I.
DR InterPro; IPR003584; HHH_2.
DR InterPro; IPR001532; XPG_I.
DR Pfam; PF01367; 5_3_exonuclease; 1.
DR Pfam; PF00867; XPG_I; 1.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS; PR00853; XPGRADSUPER.
DR SMART; SM00279; HHH2; 1.
DR SMART; SM00484; XPGI; 1.
DR SMART; SM00485; XPGN; 1.
SQ SEQUENCE 382 AA; 42865 MW; 1C648936A232D460 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVVFDGKPP 91
DB 78 GIKPVVFDGKPP 90
|||||

Query Match
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVVFDGKPP 91
DB 126 GIKPVVFDGKPP 138
|||||

RESULT 13
O93634
ID O93634 PRELIMINARY; PRT; 340 AA.
AC O93634;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ENDO/EXONUCLEASE.
GN FEN-1.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 3638;
RX MEDLINE=99415851; PubMed=10486005;

```

```

Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVVFDGKPP 91
DB 78 GIKPVVFDGKPP 90
|||||

RESULT 12
O9YFY5
ID O9YFY5 PRELIMINARY; PRT; 401 AA.
AC O9YFY5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 401AA LONG HYPOTHETICAL FLAP ENDONUCLEASE-1.
GN APE0115.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcaceae; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=993110339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,
Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000058; BAA79026.1; -.
DR HSSP; Q58839; 1A76.
DR InterPro; IPR000513; EXO_N_I.
DR InterPro; IPR003584; HHH_2.
DR InterPro; IPR001532; XPG_I.
DR Pfam; PF00867; XPG_I; 1.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS; PR00853; XPGRADSUPER.
DR SMART; SM00279; HHH2; 1.
DR SMART; SM00484; XPGI; 1.
DR SMART; SM00485; XPGN; 1.
KW Endonuclease; Complete proteome.
SQ SEQUENCE 401 AA; 45838 MW; 5ECA3DA519D75B0C CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 GIKPVVFDGKPP 91
DB 126 GIKPVVFDGKPP 138
|||||

RESULT 13
O93634
ID O93634 PRELIMINARY; PRT; 340 AA.
AC O93634;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ENDO/EXONUCLEASE.
GN FEN-1.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 3638;
RX MEDLINE=99415851; PubMed=10486005;

```

```
RA DiRuggiero J., Brown J.R., Bogert A.P., Robb F.T.;  
RT "DNA repair systems in Archaea: mementos from the last universal  
RT common ancestor?";  
RL J. Mol. Evol. 49:474-484(1999).  
DR EMBL; AF013497; AAD01514.1; -  
DR HSSP; Q58839; 1A76.  
DR InterPro; IPR000513; Exo_N_I.  
DR InterPro; IPR003584; HHH_2.  
DR InterPro; IPR001532; XPG_I.  
DR Pfam; PF00867; XPG_I; 1.  
DR Pfam; PF00752; XPG_N; 1.  
DR PRINTS; PR00853; XPGGRADSUPER.  
DR SMART; SM00279; HhH2; 1.  
DR SMART; SM00484; XPGI; 1.  
DR SMART; SM00485; XPGN; 1.  
KW Exonuclease.  
SQ SEQUENCE 340 AA; 38738 MW; D86D3D0F999E5D1E CRC64;  
  
Query Match 3.2%; Score 12; DB 1; Length 340;  
Best Local Similarity 100.0%; Pred. No. 0.0013; Mismatches 0; Indels 0; Gaps 0;  
Matches 12; Conservative 0;  
  
QY 77 EAGIRPVYVFDG 88  
DB 70 EAGIRPVYVFDG 81  
|||||  
  
RESULT 14  
Q9JHW7 PRELIMINARY; PRT; 380 AA.  
AC Q9JHW7;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE FLAG STRUCTURE-SPECIFIC ENDONUCLEASE.  
GN Fen-1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI_TaxID=10116;  
RN [1]  
RN SEQUENCE FROM N.A.  
RA Chen D., Cao G., Yang S., Li M., Chen J.;  
RT "Cloning and characterization of a rat DNA structure-specific  
RT endonuclease (Fen-1).";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF281018; AAF81265.1; -  
DR HSSP; Q58839; 1A76.  
DR InterPro; IPR002421; 5_3-exonuclease.  
DR InterPro; IPR000513; Exo_N_I.  
DR InterPro; IPR003584; HHH_2.  
DR InterPro; IPR001532; XPG_I.  
DR Pfam; PF01367; 5_3-exonuclease; 1.  
DR Pfam; PF00867; XPG_I; 1.  
DR Pfam; PF00752; XPG_N; 1.  
DR PRINTS; PR00853; XPGGRADSUPER.  
DR SMART; SM00279; HhH2; 1.  
DR SMART; SM00484; XPGI; 1.  
DR SMART; SM00485; XPGN; 1.  
DR PROSITE; PS00842; XPG_2; 1.  
KW Endonuclease.  
SQ SEQUENCE 380 AA; 42622 MW; 61698CE16F182136 CRC64;  
  
Query Match 2.9%; Score 11; DB 11; Length 380;  
Best Local Similarity 100.0%; Pred. No. 0.015;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 34 DASMSIYQPLI 44  
DB 34 DASMSIYQPLI 44  
|||||  
  
RESULT 15
```

```
Q9GZ01 PRELIMINARY; PRT; 650 AA.  
AC Q9GZ01;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE FLAP ENDONUCLEASE-1.  
GN FEN1.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI_TaxID=5833;  
RN [1]  
RN SEQUENCE FROM N.A.  
RA Casta L.J. III, Schmutte C., Taraschi T.F.;  
RT "Flap Endonuclease-1 (Plasmodium falciparum).";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF278764; AAG01445.1; -  
DR HSSP; Q58839; 1A76.  
DR InterPro; IPR002421; 5_3-exonuclease.  
DR InterPro; IPR000513; Exo_N_I.  
DR InterPro; IPR003584; HHH_2.  
DR InterPro; IPR001532; XPG_I.  
DR Pfam; PF01367; 5_3-exonuclease; 1.  
DR Pfam; PF00867; XPG_I; 1.  
DR Pfam; PF00752; XPG_N; 1.  
DR PRINTS; PR00853; XPGGRADSUPER.  
DR SMART; SM00475; 53EXOC; 1.  
DR SMART; SM00279; HhH2; 1.  
DR SMART; SM00484; XPGI; 1.  
DR SMART; SM00485; XPGN; 1.  
KW Endonuclease.  
SQ SEQUENCE 650 AA; 73989 MW; 2E9464DAB7F45B1D CRC64;  
  
Query Match 2.6%; Score 10; DB 5; Length 650;  
Best Local Similarity 100.0%; Pred. No. 0.27; Mismatches 0; Indels 0; Gaps 0;  
Matches 10; Conservative 0;  
  
QY 228 CILCGCDYCD 237  
DB 240 CILCGCDYCD 249  
|||||  
  
RESULT 16  
Q9U0K1 PRELIMINARY; PRT; 672 AA.  
AC Q9U0K1;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE PUTATIVE FLAP EXONUCLEASE.  
GN MAL4P2.21.  
OS Plasmodium falciparum (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI_TaxID=36329;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=3D7;  
RA Devlin K., Bowman S., Churcher C., Harris B., Lawson D.,  
RA Quail M., Barrell B.;  
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL035475; CAB62862.1; -  
DR HSSP; Q58839; 1A76.  
DR InterPro; IPR002421; 5_3-exonuclease.  
DR InterPro; IPR000513; Exo_N_I.  
DR InterPro; IPR003584; HHH_2.  
DR InterPro; IPR001532; XPG_I.  
DR Pfam; PF01367; 5_3-exonuclease; 1.  
DR Pfam; PF00867; XPG_I; 1.  
DR Pfam; PF00752; XPG_N; 1.  
DR PRINTS; PR00853; XPGGRADSUPER.  
DR SMART; SM00475; 53EXOC; 1.  
DR SMART; SM00279; HhH2; 1.  
DR SMART; SM00484; XPGI; 1.
```

```

DR SMART; SM00485; XPGN; 1.
KW Exonuclease.
SQ SEQUENCE 672 AA; 76680 MW; 6C2CDBB95F40042 CRC64;

Query Match      2.6%; Score 10; DB 5; Length 672;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 CILGCGDYCD 237
Db 240 CILGCGDYCD 249
|||||

RESULT 17
Q9HJDA PRELIMINARY; PRT; 336 AA.
AC Q9HJDA;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DNA REPAIR PROTEIN RAD2 RELATED PROTEIN.
GN TAL035.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaaceae;
OC Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum".
RL Nature 407:508-513(2000).
DR EMBL; AL45066; CAC12164.1; -.
DR HSSP; Q58839; IAY6.
DR InterPro; IPR000513; Exo_N_I.
DR InterPro; IPR003584; HHH_2.
DR InterPro; IPR001532; XPG_I.
DR Pfam; PF00867; XPG_I; 1.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS; PR00853; XPGRADSUPER.
DR SMART; SM00279; Hhh2; 1.
DR SMART; SM00484; XPGI; 1.
DR SMART; SM00485; XPGN; 1.
KW Complete proteome.
SQ SEQUENCE 336 AA; 38772 MW; 88EFD91DC80A78FB CRC64;

Query Match      2.4%; Score 9; DB 17; Length 336;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PVYFVDGKP 90
Db 75 PVYFVDGKP 83
|||||

RESULT 18
O94305 PRELIMINARY; PRT; 378 AA.
AC O94305;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 41.6 KDA PROTEIN.
GN SPC5E4.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F1;

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Wood V., Rajandream M.A., Barrell B.G., Rieger M.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL033406; CAA21960.1; -.
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR003089; AB_hydrolase.
DR Pfam; PF00561; abhydrolase; 1.
DR PRINTS; PR00111; ABHYDROLASE.
KW Hypothetical protein.
SQ SEQUENCE 378 AA; 41623 MW; BFB219E181E04259 CRC64;

Query Match      2.4%; Score 9; DB 3; Length 378;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 345 KPTATTSAP 353
Db 292 KPTATTSAP 300
|||||

RESULT 19
Q977Z7 PRELIMINARY; PRT; 155 AA.
AC Q977Z7;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE HYPOTHETICAL PROTEIN TVG1573729.
GN TVG1573729.
OS Thermoplasma volcanum.
OC Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaaceae;
OC Thermoplasma.
OX NCBI_TaxID=50339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GSSI / DSM 4299 / JCM 9571;
RX MEDLINE=20570466; PubMed=11121031;
RA Kawashima T., Amano N., Koike H., Yamazaki M., Kanehori K., Kawamoto T.,
RA Kawashima T., Yamamoto Y., Watanabe K., Makino S.-I., Higuchi S.,
RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;
RT "Archaeal adaptation to higher temperatures revealed by genomic
RT sequence of Thermoplasma volcanum.".
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
DR EMBL; AP000996; BAB60662.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 155 AA; 17789 MW; A0D2C6A1AAB008DA CRC64;

Query Match      2.1%; Score 8; DB 17; Length 155;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 DKDAIEKL 127
Db 135 DKDAIEKL 142
|||||

RESULT 20
Q93MR6 PRELIMINARY; PRT; 278 AA.
AC Q93MR6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE RECA (FRAGMENT).
OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F1;

```

RA Makinen J., Mertsola J., Viljanen M.K., Arvilommi H., He Q.;
 RT "Application of RecA gene sequencing for identification of Bordetella
 RT holmesii, a new species associated with pertussis-like symptoms in
 RT humans";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF399657; AAK85426.1; -;
 FT NON_TER 278
 SQ SEQUENCE 278 AA; 29669 MW; DFE070AACDB7F70B CRC64;

Query Match 2.1%; Score 8; DB 2; Length 278;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 361 DKTSKAAA 368
 Db 3 DKTSKAAA 10
 |||||

RESULT 21

ID Q93MR5 PRELIMINARY; PRT; 278 AA.
 AC Q93MR5;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE RECA (FRAGMENT).
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Bordetella.
 OX NCBI_TaxID=520;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1772;
 RA Makinen J., Mertsola J., Viljanen M.K., Arvilommi H., He Q.;
 RT "Application of RecA gene sequencing for identification of Bordetella
 RT holmesii, a new species associated with pertussis-like symptoms in
 RT humans";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF399658; AAK85427.1; -;
 FT NON_TER 278
 SQ SEQUENCE 278 AA; 29669 MW; DFE070AACDB7F70B CRC64;

Query Match 2.1%; Score 8; DB 2; Length 278;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 361 DKTSKAAA 368
 Db 3 DKTSKAAA 10
 |||||

RESULT 22

ID Q932T2 PRELIMINARY; PRT; 282 AA.
 AC Q932T2;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE RECA (FRAGMENT).
 OS Bordetella parapertussis.
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Bordetella.
 OX NCBI_TaxID=519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15311, AND SIKKALA 10;
 RA Makinen J., Mertsola J., Viljanen M.K., Arvilommi H., He Q.;
 RT "Application of RecA gene sequencing for identification of Bordetella
 RT holmesii, a new species associated with pertussis-like symptoms in
 RT humans";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF399659; AAK92001.1; -;
 FT EMBL; AF399660; AAK92002.1; -;

FT NON_TER 282
 SQ SEQUENCE 282 AA; 30097 MW; DE5A9564FFE070AA CRC64;
 Query Match 2.1%; Score 8; DB 2; Length 282;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 361 DKTSKAAA 368
 Db 3 DKTSKAAA 10
 |||||

RESULT 23

ID Q980U8 PRELIMINARY; PRT; 302 AA.
 AC Q980U8;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE DNA REPAIR ENDO/EXONUCLEASE FEN-1 (RAD2) (RAD2).
 GN RAD2.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 EX MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Awayez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Erauso G., Fletcher C., Kozera C.J., Medina N., Peng X.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Theriault C., Tolstrup N.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 DR EMBL; AE006656; AAK40525.1; -;
 DR InterPro; IPR005113; EXO_N_I.
 DR InterPro; IPR003584; HHH_2.
 DR InterPro; IPR001532; XPG_I.
 DR Pfam; PF00867; XPG_I; 1.
 DR PRINTS; PR00853; XPGRADSUPER.
 DR SMART; SM00279; Hhh2; 1.
 DR SMART; SM00484; XPGI; 1.
 DR SMART; SM00485; XPGN; 1.
 KW Exonuclease; Complete proteome.
 SQ SEQUENCE 302 AA; 34332 MW; BA0A9CCE45EA1986 CRC64;

Query Match 2.1%; Score 8; DB 17; Length 302;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 YVFDGKPP 91
 Db 30 YVFDGKPP 37
 |||||

RESULT 24

ID Q976H6 PRELIMINARY; PRT; 304 AA.
 AC Q976H6;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PUTATIVE FLAP ENDONUCLEASE.
 GN ST0210.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JCM 10545 / 7;

```

RX PubMed-11572479;
RA Kwarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankal A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain 7.";
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000981; BAB5171.1; -.
KW Endonuclease; Hypothetical protein; Complete proteome.
SQ SEQUENCE 304 AA; 34392 MW; 2275E50CA28CDB3 CRC64;

Query Match 2.1%; Score 8; DB 17; Length 304;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 YVFDGKPP 91
DB 30 YVFDGKPP 37

RESULT 25
OS0123 PRELIMINARY; PRT; 343 AA.
ID OS0123
AC OS0123;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 343AA LONG HYPOTHETICAL 5' NUCLEASE.
GN PH1415.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN SEQUENCE FROM N.A.
RP STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kwarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Koshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL; AP000006; BAA30521.1; -.
DR HSSP; O58839; 1A76.
DR InterPro; IPR000513; EXO_N1.
DR InterPro; IPR003584; HHH_2.
DR InterPro; IPR001532; XPG_I.
DR Pfam; PF00867; XPG_I; 1.
DR Pfam; PF00752; XPG_N; 1.
DR PRINTS; PR00853; XPGRADSUPER.
DR SMART; SM00279; HH2; 1.
DR SMART; SM00484; XPGT; 1.
DR SMART; SM00485; XPGN; 1.
KW Complete proteome.
SQ SEQUENCE 343 AA; 38947 MW; 8BE0025F372C3138 CRC64;

Query Match 2.1%; Score 8; DB 17; Length 343;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 YVFDGKPP 91
DB 77 YVFDGKPP 84

RESULT 26
Q96UV5 PRELIMINARY; PRT; 354 AA.
ID Q96UV5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GLUTAMINE SYNTHETASE.
OS Hebeloma cylindrosporum.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Cortinariaceae; Hebeloma.
OX NCBI_TaxID=76867;
RN SEQUENCE FROM N.A.
RP STRAIN=H1;
RA Rodriguez-Pastrana B., Javelle A., Belleville R., Morel M., Botton B.,
RA Jacob C., Chalot M., Brun A.;
RT "Nucleotide sequence and expression of NADP-GDH and GS from the
RT ectomycorrhizal fungus Hebeloma cylindrosporum.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF411820; AAK96111.1; -.
SQ SEQUENCE 354 AA; 39225 MW; DEFEF69A36322241 CRC64;

Query Match 2.1%; Score 8; DB 3; Length 354;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
DB 268 AIEKLSKR 275

RESULT 27
Q9EQP8 PRELIMINARY; PRT; 373 AA.
ID Q9EQP8;
AC Q9EQP8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GLUTAMINE SYNTHETASE (EC 6.3.1.2).
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN SEQUENCE FROM N.A.
RP Tong Y., Wang H.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) = ADP + GLUTAMINE +
CC ORTHOPHOSPHATE.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
DR EMBL; AF150961; AAG43362.1; -.
DR InterPro; IPR001691; GLN_synth.
DR Pfam; PF00120; gln-synt; 1.
DR PROSITE; PS00180; GLNA_1; 1.
KW Ligase.
SQ SEQUENCE 373 AA; 42320 MW; 611D58CE20FB16CF CRC64;

Query Match 2.1%; Score 8; DB 11; Length 373;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
DB 273 AIEKLSKR 280

RESULT 28
Q96770 PRELIMINARY; PRT; 399 AA.
ID Q96770;
AC Q96770;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

```

```

DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE GLUTAMINE SYNTHETASE (EC 6.3.1.2) (GLUTAMATE--AMMONIA LIGASE).
GN GSI OR CG2718.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Glover D.M.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) -> ADP + GLUTAMINE +
CC ORTHOPHOSPHATE.
CC -!- SUBUNIT: HOMOCYTAMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
DR EMBL; AJ012460; CAAL0031.1; -.
DR FlyBase; FBgn0001142; Gsl.
DR InterPro; IPR001691; GLN_synth.
DR Pfam; PF00120; gln-synt; 1.
DR PROSITE; PS00180; GLNA_1; 1.
DR PROSITE; PS00181; GLNA-ATP; 1.
KW Ligase; Nitrogen fixation.
KW Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
SQ SEQUENCE 399 AA; 44481 MW; 57EALB03348244F3 CRC64;

Query Match 2.1%; Score 8; DB 5; Length 399;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
DB 309 AIEKLSKR 316
|||||||

RESULT 29
Q9VPKO PRELIMINARY; PRT; 399 AA.
ID Q9VPKO
AC Q9VPKO;
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Feariera S., Fleischmann W.,
RA Foster C., Gargiilan A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```

```

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Y, CN BW SP;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mangall C.J.,
RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) -> ADP + GLUTAMINE +
CC ORTHOPHOSPHATE.
CC -!- SUBUNIT: HOMOCYTAMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
DR EMBL; AE003590; AAF51547.1; -.
DR EMBL; AY058730; AAL13959.1; -.
DR FlyBase; FBgn0001142; Gsl.
DR InterPro; IPR001691; GLN_synth.
DR Pfam; PF00120; gln-synt; 1.
DR PROSITE; PS00180; GLNA_1; 1.
DR PROSITE; PS00181; GLNA-ATP; 1.
KW Ligase.
SQ SEQUENCE 399 AA; 44396 MW; 935E8D1D9927ACCC CRC64;

Query Match 2.1%; Score 8; DB 5; Length 399;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AIEKLSKR 130
DB 309 AIEKLSKR 316
|||||||

RESULT 30
Q9AWB2 PRELIMINARY; PRT; 573 AA.
ID Q9AWB2
AC Q9AWB2;
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Feariera S., Fleischmann W.,
RA Foster C., Gargiilan A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

```


DR PROSITE; PS00365; NIR_SIR; 1.
 KW Transit peptide; Chloroplast.
 FT TRANSIT 1 55 CHLOROPLAST.
 FT CHAIN 56 >573 FERREDOXIN:SULFITE REDUCTASE.
 FT NON_TER 573 573
 SQ SEQUENCE 573 AA; 63821 MW; FD02A907B746B2D9 CRC64;

Query Match 2.1%; Score 8; DB 10; Length 573;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 348 ATTSAPLK 355
 Db 8 ATTSAPLK 15

RESULT 31

Q9LDM8 PRELIMINARY; PRT; 770 AA.
 AC Q9LDM8;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PUTATIVE SUBTILISIN PRECURSOR (PUTATIVE PRE-PRO-SUBTILISIN PRECURSOR)
 DE (SUBTILISIN-TYPE PROTEASE PRECURSOR).
 GN SCS1 OR SSTP-1.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. MAPLE PRESTO; TISSUE=SEED COAT;
 RA Batchelor A.K., Boutiller K., Miller S., Labbe H., Bowman L., Hu M.,
 RA Johnson D.A., Gijzen M., Miki B.L.;
 RT "The seed-coat specific expression of a subtilisin gene, SCS1, from
 RT soybean.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. HAROVINTON;
 RA Batchelor A.K., Boutiller K., Miller S., Labbe H., Bowman L., Hu M.,
 RA Johnson D.A., Gijzen M., Miki B.L.;
 RT "The seed-coat specific expression of a subtilisin gene, SCS1, from
 RT soybean.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN=CV. RESNIK;
 RA Beilinson V., Moskalenko O., Livingston D.S. III, Reverdatto S.V.,
 RA Jung R., Nielsen N.C.;
 RT "Subtilisin-type protease from soybean seeds.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ276710; CAB87247.1; -;
 DR EMBL; AJ276407; CAB87246.1; -;
 DR EMBL; AF160513; AAG38994.1; -;
 DR HSSP; P27693; 1AH2.
 DR MEROPS; S08.UPA; -;
 DR InterPro; IPR003137; PA.
 DR Fram; PF02225; PA; 1.
 DR Pfam; PF00082; Peptidase_S8; 2.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Signal; Protease.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 117 770 PUTATIVE SUBTILISIN PRECURSOR.
 SQ SEQUENCE 770 AA; 82693 MW; 7DB833C5A16C8037 CRC64;

Query Match 2.1%; Score 8; DB 10; Length 770;
 Best Local Similarity 100.0%; Pred. No. 37;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 152 PVVEAPSE 159
 Db 707 PVVEAPSE 714

RESULT 32

Q98CQ1 PRELIMINARY; PRT; 802 AA.
 AC Q98CQ1;
 DT 01-OCT-2001 (Tremblrel. 18, Created)
 DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
 DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
 DE PHENYLALANINE-TRNA LIGASE BETA CHAIN (EC 6.1.1.20).
 GN MLL5053.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP003005; BABSL570.1; -;
 DR InterPro; IPR002547; trna_bind.
 DR Pfam; PF01588; trna_bind; 2.
 KW Ligase; Complete proteome.
 SQ SEQUENCE 802 AA; 85169 MW; 32C01501A34D5241 CRC64;

Query Match 2.1%; Score 8; DB 16; Length 802;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 GGQTALKL 250
 Db 526 GGQTALKL 533

RESULT 33

Q27446 PRELIMINARY; PRT; 862 AA.
 AC Q27446;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE RNA POLYMERASE II, SECOND LARGEST SUBUNIT (FRAGMENT).
 OS Artemia salina (Brine shrimp).
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
 OC Artemiidae; Artemia.
 OX NCBI_TaxID=85549;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95041334; PubMed=7953533;
 RA Sidow A., Thomas W.K.;
 RT "A molecular evolutionary framework for eukaryotic model organisms.";
 RL Curr. Biol. 4:596-603(1994).
 CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE +
 CC RNA(N).
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
 DR EMBL; U10332; AAC83398.1; -;
 DR InterPro; IPR001572; RNA_pol_B.
 DR Pfam; PF00562; RNA_pol_B; 1.
 DR PROSITE; PS01166; RNA_POL_BETA; 1.
 KW DNA-directed RNA polymerase; Transcription; Transferase.

```
FT NON_TER 1
FT NON_TER 862 862
SQ SEQUENCE 862 AA; 97356 MW; CA2D0FB15F9E6296 CRC64;

Query Match 2-18; Score 8; DB 5; Length 862;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 AKRYSKRD 106
Db 694 AKRYSKRD 701
|||||

RESULT 34
Q98FF9 PRELIMINARY; PRT; 59 AA.
AC Q98FF9;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE MSL3793 PROTEIN.
GN MSL3793.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003002; BAB50608.1; -.
KW Complete proteome.
SQ SEQUENCE 59 AA; 6785 MW; F804C038B6AFBF23 CRC64;

Query Match 1-8; Score 7; DB 16; Length 59;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 KDAIEKL 127
Db 46 KDAIEKL 52
|||||

RESULT 35
Q9JYG3 PRELIMINARY; PRT; 63 AA.
AC Q9JYG3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN NMB1598.
GN NMB1598.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MC58; SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tetelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
```

```
RA Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815(2000).
DR EMBL; AE002510; AAP41951.1; -.
DR TIGR; NMB1598; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 63 AA; 7758 MW; 2C36B20E0D402666 CRC64;

Query Match 1-8; Score 7; DB 16; Length 63;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 KFSYFG 28
Db 4 KFSYFG 10
|||||

RESULT 36
O32527 PRELIMINARY; PRT; 65 AA.
ID O32527
AC O32527;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE HYPOTHETICAL 7.5 KDA PROTEIN (FRAGMENT).
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K-12;
RX MEDLINE=95189796; PubMed=7883769;
RA Kikuda H., Hosono K., Kazuya S., Ichihara S.;
RT "Identification and characterization of the ackA (acetate kinase A)-
RT pta (phosphotransacetylase) operon and complementation analysis of
RT acetate utilization by an ackA-pta deletion mutant of Escherichia
RT coli.";
RL J. Biochem. 116:916-922(1994).
DR EMBL; D17576; BAA20540.1; -.
KW Hypothetical protein.
FT NON_TER 65
SQ SEQUENCE 65 AA; 7507 MW; E7747C36ELC534B6 CRC64;

Query Match 1-8; Score 7; DB 2; Length 65;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 54 LTNEAGE 60
Db 43 LTNEAGE 49
|||||

RESULT 37
Q9ZWZ6 PRELIMINARY; PRT; 65 AA.
ID Q9ZWZ6
AC Q9ZWZ6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE GP84.
OS Mycobacteriophage TM4.
OC Viruses.
OX NCBI_TaxID=88870;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20110038; PubMed=10645443;
RA Ford M.E., Stenstrom C., Hendrix R.W., Hatfull G.F.;
RT "Mycobacteriophage TM4: Genome structure and gene expression.";
RL Tuberc. Lung Dis. 79:63-73(1998).
RN [2]
```

RP SEQUENCE FROM N.A.
 RA Ford M.E., Stenstrom C., Hendrix R.W., Hatfull G.F.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RM EMBL; AF068845; AAD17649.1; -
 SQ SEQUENCE 65 AA; 7613 MW; B9BD107B15F00AEC CRC64;

Query Match 1.8%; Score 7; DB 9; Length 65;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 VEAPSEA 160
 |||||
 Db 22 VEAPSEA 28

RESULT 38
 Q9K880 PRELIMINARY; PRT; 67 AA.
 ID Q9K880
 AC Q9K880
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE BH3126 PROTEIN.
 GN BH3126.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirano C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis."
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL; AP001517; BAB06845.1; -
 KW Complete proteome.
 SQ SEQUENCE 67 AA; 7735 MW; 0825029A4C8FBC4E CRC64;

Query Match 1.8%; Score 7; DB 16; Length 67;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 KAEKIK 329
 |||||
 Db 14 KAEKIK 20

RESULT 39
 Q9QJD7 PRELIMINARY; PRT; 68 AA.
 ID Q9QJD7
 AC Q9QJD7
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ENVELOPE PROTEIN (FRAGMENT).
 GN ENV.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=132N;
 RX MEDLINE=20170331; PubMed=10708058;
 RA Casado C., Urtasun I., Martin-Walther M.V., Garcia S., Rodriguez C.,
 RA del Romero J., Lopez-Galindez C.;
 RT "Genetic analysis of HIV-1 samples from Spain."
 RL J. Acquir. Immune Defic. Syndr. 23:68-74(2000).
 DR EMBL; AF152825; AAF08463.1; -
 DR InterPro; IPR000777; GP120.

DR Pfam; PF00516; GP120; 1.
 KW AIDS; Coat protein; Glycoprotein.
 FT NON_TER 1
 FT NON_TER 68
 SQ SEQUENCE 68 AA; 7629 MW; 303DE17F719FD29E CRC64;

Query Match 1.8%; Score 7; DB 15; Length 68;
 Best Local Similarity 100.0%; Pred. No. 43;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 LTADEV 118
 |||||
 Db 6 LTADEV 12

RESULT 40
 Q37945 PRELIMINARY; PRT; 78 AA.
 ID Q37945
 AC Q37945
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE SIMILARITY TO GENE 24 OF PHAGE P22.
 GN 24.
 OS Bacteriophage L.
 OC Viruses.
 OX NCBI_TaxID=45441;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97444294; PubMed=9300826;
 RA Schickmaier P., Schmieger H.;
 RT "Sequence comparison of the genes for immunity, DNA replication, and
 RT cell lysis of the P22-related Salmonella phages ES18 and L."
 RL Gene 195:93-100(1997).
 DR EMBL; X94331; CAA63998.1; -
 DR HSP; P04891; IAT.
 SQ SEQUENCE 78 AA; 8856 MW; 232A83994EB77763 CRC64;

Query Match 1.8%; Score 7; DB 9; Length 78;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 DRVTKAI 325
 |||||
 Db 58 DRVTKAI 64

RESULT 41
 Q9P192 PRELIMINARY; PRT; 96 AA.
 ID Q9P192
 AC Q9P192
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE PRO1843.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
 RA He F.;
 RT "Functional prediction of the coding sequences of 79 new genes deduced
 RT by analysis of cDNA clones from human fetal liver."
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF119854; AAF69608.1; -
 SQ SEQUENCE 96 AA; 10994 MW; 69EE9D4A1DDAC566 CRC64;

Query Match 1.8%; Score 7; DB 4; Length 96;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 126 KLSKRTV 132
Db 6 KLSKRTV 12

RESULT 42
Q9I449
ID Q9I449 PRELIMINARY; PRT; 97 AA.
AC Q9I449;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN PAI295.
GN PAI295.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PAOI;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "complete genome sequence of Pseudomonas aeruginosa PAOI, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
RW EMBL; AE004559; AAG04684.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 97 AA; 11361 MW; 3BE50976A64485D2 CRC64;

Query Match 1.8%; Score 7; DB 16; Length 97;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 DVAKVLE 215
Db 58 DVAKVLE 64

RESULT 43
Q9YCG8
ID Q9YCG8 PRELIMINARY; PRT; 103 AA.
AC Q9YCG8;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE HYPOTHETICAL 11.1 KDA PROTEIN APE1289.
GN APE1289.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K1;
RX MEDLINE=99510339; PubMed=10382966;
RA Kavarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaï A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1."
RL DNA Res. 6:83-101(1999).
RW EMBL; AF000061; BAA80280.1; -.
KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 118 AA; 14091 MW; C25D77060E06357D CRC64;

Query Match 1.8%; Score 7; DB 10; Length 118;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 VYVFDGK 89
Db 74 VYVFDGK 80

RESULT 45
Q9SU26
ID Q9SU26 PRELIMINARY; PRT; 128 AA.
AC Q9SU26;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RIBOSOMAL PROTEIN L7AE-LIKE (AT4G13600/TLP17_190).
GN TLP17_190 OR AT4G13600.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
RA Bancroft I., Meeses H.W., Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

```

```

RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Banh J., Bowser L.,
RA Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y., Ishida J.,
RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M.,
RA Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M., Seki M.,
RA Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K., Davis R.W.,
RA Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Koesema E., Chen H., Cheuk R., Kim C.J., Meyers M.C., Shinn P.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Ohodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL049730; CAB53753.1; -.
DR EMBL; AL161534; CAB78303.1; -.
DR EMBL; AF361618; AAK32786.1; -.
DR EMBL; AY055095; AAL05895.1; -.
DR HSP; P55769; 1E7K.
DR InterPro; IPR004037; Ribosomal_L7Ae.
DR InterPro; IPR004038; Ribosomal_L7Ae_L30e_S12e.
DR Pfam; PF01248; Ribosomal_L7Ae; 1.
DR PRINTS; PR00881; L7ARS6FAMILY.
DR PROSITE; PS01082; RIBOSOMAL_L7AE; UNKNOWN_1.
SQ SEQUENCE 128 AA; 13999 MW; 3AA520DD6187C96A CRC64;

```

```

Query Match      1.8%; Score 7; DB 10; Length 128;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 121 KDAIEKL 127
DB 120 KDAIEKL 126

```

Search completed: November 5, 2002, 16:25:00
Job time : 76 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 11:03:23 ; Search time 1751.33 Seconds
(without alignments)
16501.458 Million cell updates/sec

Title: US-09-805-311-5

Perfect score: 1381
Sequence: 1 cgcaccacgcgtccggccac.....ttgaaaaaaaaaaaaaaaa 1381

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_em.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
------------	-------	-------	--------	-------	-------------

1	1381	100.0	1381	6	AR152405	AR152405 Sequence
2	1267	91.7	1463	6	AR152403	AR152403 Sequence
3	1216	88.1	1541	6	AR152404	AR152404 Sequence
4	850	61.5	1478	6	AR152406	AR152406 Sequence
5	58	4.2	1354	8	AB021666	AB021666 Oryza sat
6	58	4.2	115907	8	AC104713	AC104713 Oryza sat
7	30	2.2	110811	8	AF0004520	AF0004520 Lotus jap
8	27	2.0	161494	2	AF215846	AF215846 Homo sapi
9	27	2.0	220761	2	AL670227	AL670227 Mus muscu
10	26	1.9	1203	3	AF100286	AF100286 Trypanoso
11	26	1.9	1395	3	AF106931	AF106931 Medicago
12	26	1.9	192363	10	AC068906	AC068906 Mus muscu
13	26	1.9	193026	2	AC108055	AC108055 Homo sapi
14	26	1.9	196367	9	AC020740	AC020740 Homo sapi
15	26	1.9	197515	2	AC087134	AC087134 Mus muscu
16	26	1.9	227764	10	AC034109	AC034109 Mus muscu
17	26	1.9	234701	2	AC025116	AC025116 Mus muscu
18	25	1.8	1232	9	BSA420445	AJ420445 Homo sapi
19	25	1.8	1808	9	BC003123	BC003123 Homo sapi
20	25	1.8	1834	9	BC017335	BC017335 Homo sapi
21	25	1.8	2830	9	BSM80117	AL117584 Homo sapi
22	25	1.8	90141	9	AC035150	AC035150 Homo sapi
23	25	1.8	109088	9	AL137011	AL137011 Human DNA
24	25	1.8	148210	2	AC098661	AC098661 Rattus no
25	25	1.8	160434	2	AC024144	AC024144 Mus muscu
26	25	1.8	160492	9	AP001884	AP001884 Homo sapi
27	25	1.8	162978	9	AC021070	AC021070 Homo sapi
28	25	1.8	164382	2	AC018421	AC018421 Homo sapi
29	25	1.8	165203	9	AL589693	AL589693 Human DNA
30	25	1.8	170896	2	AC103898	AC103898 Bos tauru
31	25	1.8	176150	2	AP001098	AP001098 Homo sapi
32	25	1.8	176292	9	AP002407	AP002407 Homo sapi
33	25	1.8	178726	2	AC098950	AC098950 Rattus no
34	25	1.8	182365	2	AC105354	AC105354 Rattus no
35	25	1.8	185721	2	AC019108	AC019108 Homo sapi
36	25	1.8	267156	6	AX336388	AX336388 Sequence
37	25	1.8	267156	9	U66059	U66059 Human getml
38	24	1.7	230	8	AF124738	AF124738 Zea mays
39	24	1.7	292	6	AX314458	AX314458 Sequence
40	24	1.7	387	6	AX069660	AX069660 Sequence
41	24	1.7	573	6	AX186327	AX186327 Sequence
42	24	1.7	597	6	AX187052	AX187052 Sequence
43	24	1.7	648	10	S82852	S82852 tmk-thymidi
44	24	1.7	700	3	DDIUPDGP	M27639 Dictyostell
45	24	1.7	766	9	AF339830	AF339830 Homo sapi

ALIGNMENTS

RESULT 1	AR152405	AR152405	1381 bp	DNA	linear	PAT 08-AUG-2001
LOCUS	Sequence	5 from patent US 6232527.				
DEFINITION	AR152405					
ACCESSION	AR152405					
VERSION	AR152405.1	GI:15118455				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 1381)					
AUTHORS	Mahajan,P.B.					
TITLE	Maize Rad2/FEN-1 orthologues and uses thereof					
JOURNAL	Patent: US 6232527-A 5 15-MAY-2001;					
FEATURES	Location/Qualifiers					
source	1..1381					
BASE COUNT	441 a	269 c	346 g	325 t		
ORIGIN						
Query Match	100.0%	Score 1381;	DB 6;	Length 1381;		
Best Local Similarity	100.0%;	Pred. No. 0;				
Matches 1381;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		

QY	1	CGACCCACGCTCCGGCCACAGCCGCGCAGACGAGATGGGCATCAAGGGTTTGACGAAA	60
Db	1	CGACCCACGCTCCGGCCACAGCCGCGCAGACGAGATGGGCATCAAGGGTTTGACGAAA	60
QY	61	CTGCTGGCGGCAATAGTCGGCCCAAGGGATGAAGGACGAGAAGTTTCAGAGCTACTTCGGC	120
Db	61	CTGCTGGCGGCAATAGTCGGCCCAAGGGATGAAGGACGAGAAGTTTCAGAGCTACTTCGGC	120
QY	121	CGCAAAATCGCGTCGACGCCAGCATGACATATACCAAGTTCCCTGATTTAGTTGGAAGG	180
Db	121	CGCAAAATCGCGTCGACGCCAGCATGACATATACCAAGTTCCCTGATTTAGTTGGAAGG	180
QY	181	ACAGGATGGAATCTCTCAAAATGAAGCTGGTGAAGTCACTAGTCATTTGCAAGGAATG	240
Db	181	ACAGGATGGAATCTCTCAAAATGAAGCTGGTGAAGTCACTAGTCATTTGCAAGGAATG	240
QY	241	TTCAACCGGCAATAGATTACTGGAGCGGGAATCAAGCCAGTTTATCTTTTGGATGCG	300
Db	241	TTCAACCGGCAATAGATTACTGGAGCGGGAATCAAGCCAGTTTATCTTTTGGATGCG	300
QY	301	AAGCCTCTGATATGAAGAACAAAGAGCTTGCTTAAAGATACTCAAAAAGAGATGATCA	360
Db	301	AAGCCTCTGATATGAAGAACAAAGAGCTTGCTTAAAGATACTCAAAAAGAGATGATCA	360
QY	361	ACCAAGATCTGACTGAGGCAGTAGAGGTAGGAGATAAAGATCGGATTTGAAAATTTGAGC	420
Db	361	ACCAAGATCTGACTGAGGCAGTAGAGGTAGGAGATAAAGATCGGATTTGAAAATTTGAGC	420
QY	421	AAGAGGACTCTAAAGGTCACAAAGCAACACACAGAGATTGTAAACGGCTATTAAAGACTT	480
Db	421	AAGAGGACTCTAAAGGTCACAAAGCAACACAGAGATTGTAAACGGCTATTAAAGACTT	480
QY	481	ATGGGGTTCCTGTTAGAGGACCTTCTGAAGCAGACCAAAATGTCGAGCCCTTTGC	540
Db	481	ATGGGGTTCCTGTTAGAGGACCTTCTGAAGCAGACCAAAATGTCGAGCCCTTTGC	540
QY	541	ATAAACCAGTAAGGTGTCGCTTGCTTCAGAAAGATGGACTCCCTTACTTTTGGGGCT	600
Db	541	ATAAACCAGTAAGGTGTCGCTTGCTTCAGAAAGATGGACTCCCTTACTTTTGGGGCT	600
QY	601	CCACGGTTCCTGTCATTTAATGGATCCAAGTTCGAAGTCCAAAGATACCTGTGATGGAATTT	660
Db	601	CCACGGTTCCTGTCATTTAATGGATCCAAGTTCGAAGTCCAAAGATACCTGTGATGGAATTT	660
QY	661	GATGTTGCCAAGGTTTGGAGGAGCTTGAACCTACCAATGGACCAAGTTCAATTTGTCG	720
Db	661	GATGTTGCCAAGGTTTGGAGGAGCTTGAACCTACCAATGGACCAAGTTCAATTTGTCG	720
QY	721	ATCCTGTGTGATGTGACTATTGTGATAGCATCAAAGGTATCGGGGGGCAAAACAGCTCTG	780
Db	721	ATCCTGTGTGATGTGACTATTGTGATAGCATCAAAGGTATCGGGGGGCAAAACAGCTCTG	780
QY	781	AAACTTATCTCAACATGGGTCCATAGAAGCATCTTGAGATCTTATAAAGACAGA	840
Db	781	AAACTTATCTCAACATGGGTCCATAGAAGCATCTTGAGATCTTATAAAGACAGA	840
QY	841	TATCAAAATCTCTGAGGACTGCGCTTTACCAAGAGCTCGACGCTTTGTTCAAGGAGCCCTAAT	900
Db	841	TATCAAAATCTCTGAGGACTGCGCTTTACCAAGAGCTCGACGCTTTGTTCAAGGAGCCCTAAT	900
QY	901	GTCAATTTGGATATTCCTGAOCTAAATPGACTGCACCTGATGAGGAGGTCCTCATAGT	960
Db	901	GTCAATTTGGATATTCCTGAOCTAAATPGACTGCACCTGATGAGGAGGTCCTCATAGT	960
QY	961	TTCCGTTAAAGATAATGGTTTCAACGAAGATCGGGTGACAAAGGCCATAGAGAAGATC	1020
Db	961	TTCCGTTAAAGATAATGGTTTCAACGAAGATCGGGTGACAAAGGCCATAGAGAAGATC	1020
QY	1021	AAATCTGCCAAGATAAATTCGTCGAAGGAAGACTTCGAGTCTCTTTTCAAGCCCACTGCC	1080
Db	1021	AAATCTGCCAAGATAAATTCGTCGAAGGAAGACTTCGAGTCTCTTTTCAAGCCCACTGCC	1080
QY	1081	ACCACATCAGCACCGCTAAACCGAAGGAGACTTCGATATAAAACAAAGCAGGCTGCG	1140
Db	1081	ACCACATCAGCACCGCTAAACCGAAGGAGACTTCGATATAAAACAAAGCAGGCTGCG	1140
QY	1141	AACAAGAAAAACAAGGCTGCTGGAAAAAGAAATAATCTTTGGATGCTTGATGTACAACTA	1200
Db	1141	AACAAGAAAAACAAGGCTGCTGGAAAAAGAAATAATCTTTGGATGCTTGATGTACAACTA	1200
QY	1201	CGACTACGAAGACGCGTGGCTGATCACTTCGCTTAGATTATTTAACTCCCTGTTTAA	1260
Db	1201	CGACTACGAAGACGCGTGGCTGATCACTTCGCTTAGATTATTTAACTCCCTGTTTAA	1260
QY	1261	ACTCAGAGCTTTGTGTAAGAGTTCCGCCCATGTTTCAAGCTGGGTAACTTAACTTGTGTTG	1320
Db	1261	ACTCAGAGCTTTGTGTAAGAGTTCCGCCCATGTTTCAAGCTGGGTAACTTAACTTGTGTTG	1320
QY	1321	AAGAGATTGGTGTACCAAGTAACAAAACCTTATCGCTGTTTTTTCAGAAAAAAGAAAAA	1380
Db	1321	AAGAGATTGGTGTACCAAGTAACAAAACCTTATCGCTGTTTTTTCAGAAAAAAGAAAAA	1380
QY	1381	A	1381
Db	1381	A	1381
RESULT 2			
LOCUS AR152403 1463 bp DNA linear PAT 08-AUG-2001			
DEFINITION Sequence 1 from patent US 6232527.			
ACCESSION AR152403			
VERSION AR152403.1 GI:15118453			
KEYWORDS			
SOURCE Unknown.			
ORGANISM Unclassified.			
REFERENCE 1 (bases 1 to 1463)			
AUTHORS Mahajan,P.,B.			
TITLE Maize Rad2/FEN-1 orthologues and uses thereof			
JOURNAL Patent: US 6232527-A 1 15-MAY-2001;			
FEATURES Location/Qualifiers			
1..1463			
Source /organism="unknown"			
BASE COUNT 466 a 292 c 361 g 344 t			
ORIGIN			
Query Match 91.7%; Score 1267; DB 6; Length 1463;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	16	GCCACAGCCCGCCGACGACGAGATGGCATCAAGGGTTTGACGAAAACCTGCTGGCGGACAAT	75
Db	64	GCCACAGCCCGCCGACGACGAGATGGCATCAAGGGTTTGACGAAAACCTGCTGGCGGACAAT	123
QY	76	GGCCCCAAGCGCATGAAGGACGAGAGTTTCGAGAGCTACTTGGCGCGCAAAATCGCCGTC	135
Db	124	GGCCCCAAGCGCATGAAGGACGAGAGTTTCGAGAGCTACTTGGCGCGCAAAATCGCCGTC	183
QY	136	GAGCCGACGATGAGCATATACCAGTTCCCTGATTTAGTTTGAAGGACGACGCGATGCAAACT	195
Db	184	GAGCCGACGATGAGCATATACCAGTTCCCTGATTTAGTTTGAAGGACGACGCGATGCAAACT	243
QY	196	CTCACAAATGAAGCTGGTGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACAATA	255
Db	244	CTCACAAATGAAGCTGGTGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACAATA	303
QY	256	AGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTTTGTATGGCANGCCTCCTGATATG	315
Db	304	AGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTTTGTATGGCANGCCTCCTGATATG	363
QY	316	AAGAAACAGAGCTTCTGCTAAAAGATACTCAAAAAGAGATGATGCAACCAAGAGATCTGACT	375
Db	364	AAGAAACAGAGCTTCTGCTAAAAGATACTCAAAAAGAGATGATGCAACCAAGAGATCTGACT	423

QY 376 GAGGCAGTAGAGTAGAGATAAAGATCGGATTCAAAAATTGAGCAAGAGGACTGTAAAG 435
|||||
Db 424 GAGGCAGTAGAGTAGAGATAAAGATCGGATTCAAAAATTGAGCAAGAGGACTGTAAAG 483
QY 436 GTCACAAGGCACACACAGAGATGTTAAACGGCTATTAGACTTATGCGGGTTCCTGTT 495
|||||
Db 484 GTCACAAGGCACACACAGAGATGTTAAACGGCTATTAGACTTATGCGGGTTCCTGTT 543
QY 496 GTAGAGGCACCTTCTGAGCAGAGACAGAAATGTGACGCCCTTTGCAATAAACGATAAGGTG 555
|||||
Db 544 GTAGAGGCACCTTCTGAGCAGAGACAGAAATGTGACGCCCTTTGCAATAAACGATAAGGTG 603
QY 556 TTCGCTGTGCTTCAGAGATATGGACTCCCTTACTTTTGGGGCTCCAGCGTTCTCTGTT 615
|||||
Db 604 TTCGCTGTGCTTCAGAGATATGGACTCCCTTACTTTTGGGGCTCCAGCGTTCTCTGTT 663
QY 616 CATTTAATGGATCCCAAGTTCACAAGAAATACCTGTGATGGAATTTGATGTTGCCAAGGTT 675
Db 664 CATTTAATGGATCCCAAGTTCACAAGAAATACCTGTGATGGAATTTGATGTTGCCAAGGTT 723
QY 676 TTGGAGGAGCTTGAACCTCACCATGGACAGTTCATTGATTTGTGCATCTCTGTGTGGATGT 735
Db 724 TTGGAGGAGCTTGAACCTCACCATGGACAGTTCATTGATTTGTGCATCTCTGTGTGGATGT 783
QY 736 GACTATTGTGTAGCATCAAAAGGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCAA 795
Db 784 GACTATTGTGTAGCATCAAAAGGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCAA 843
QY 796 CATGGTCCNTAGANAGCATCTTGGGAATCTTAATAAGACAGATATCAAAATTCCTGAG 855
Db 844 CATGGTCCNTAGANAGCATCTTGGGAATCTTAATAAGACAGATATCAAAATTCCTGAG 903
QY 856 GACTGGCTTACCAAGAAAGCTCGACGCTTGTTCAGAGAGCCTAATCTCAATTTGGATAT 915
Db 904 GACTGGCTTACCAAGAAAGCTCGACGCTTGTTCAGAGAGCCTAATCTCAATTTGGATAT 963
QY 916 CTTGAGCTAAATGACTGACCTTGATCAGAGGGTCTCATAGTTTCTGGTAAAGAT 975
Db 964 CTTGAGCTAAATGACTGACCTTGATCAGAGGGTCTCATAGTTTCTGGTAAAGAT 1023
QY 976 AATGTTTTCAAGAGATCGGGTGACAAAGGCCATAGAGAAGTCAAAATCTGCCAAGAT 1035
Db 1024 AATGTTTTCAAGAGATCGGGTGACAAAGGCCATAGAGAAGTCAAAATCTGCCAAGAT 1083
QY 1036 AATGCTGCGAAGAGACTCGAGTCTCTTTTCAAGCCAACTCCACACATCAGACCG 1095
Db 1084 AATGCTGCGAAGAGACTCGAGTCTCTTTTCAAGCCAACTCCACACATCAGACCG 1143
QY 1096 CTAAACGSAAGGAGACTTCGGATAAAACAAAGCAGCTCGGAACAAAGAAACAAAG 1155
Db 1144 CTAAACGSAAGGAGACTTCGGATAAAACAAAGCAGCTCGGAACAAAGAAACAAAG 1203
QY 1156 GCTGTGGAAGAGAAATAATCTTGATGCTGTGATGATGATGATGATGATGATGATGATG 1215
Db 1204 GCTGTGGAAGAGAAATAATCTTGATGCTGTGATGATGATGATGATGATGATGATGATG 1263
QY 1216 CGTGCGCTGTGATCAGTCTCGCTTAGATTATTAACCTCCCTTTTAACTCAGAGCTTTGGT 1275
Db 1264 CGTGCGCTGTGATCAGTCTCGCTTAGATTATTAACCTCCCTTTTAACTCAGAGCTTTGGT 1323
QY 1276 AAAAGTT 1282
|||||
Db 1324 AAAAGTT 1330

RESULT 3
AR152404
LOCUS AR152404 1541 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 3 from patent US 6232527.
ACCESSION AR152404
VERSION AR152404.1 GI:15118454
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1541)
AUTHORS Mahajan,P.B.
TITLE Maize Rad2/FEN-1 orthologues and uses thereof
JOURNAL Patent: US 6232527-A 3 15-MAY-2001;
FEATURES Location/Qualifiers
1..1541
source /organism="unknown"
BASE COUNT 473 a 308 c 377 g 383 t
ORIGIN

Query Match 88.1%; Score 1216; DB 6; Length 1541;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1266; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 16 GCCACAGCCGCCGACAGAGATGGGCATCAAGGTTTGACGAAACTGCTGGCGGACAAT 75
Db 58 GCCACAGCCGCCGACAGAGATGGGCATCAAGGTTTGACGAAACTGCTGGCGGACAAT 117
QY 76 GCGCCCAAGCGGATGAAGAGCAGAAAGTTCGAGAGCTACTTCGGCGCAAAAATCGCGGTC 135
Db 118 GCGCCCAAGCGGATGAAGAGCAGAAAGTTCGAGAGCTACTTCGGCGCAAAAATCGCGGTC 177
QY 136 GAGCCAGCATCAGCATATACGAGTTCCTGATTTGAGTTGCAAGGACAGCATGGAAC 195
Db 178 GAGCCAGCATCAGCATATACGAGTTCCTGATTTGAGTTGCAAGGACAGCATGGAAC 237
QY 196 CTCACAAATGAAGCTGGTGAAGTCACTAGTCACTTTGCAAGGAATGTTCAACCGGACAATA 255
Db 238 CTCACAAATGAAGCTGGTGAAGTCACTAGTCACTTTGCAAGGAATGTTCAACCGGACAATA 297
QY 256 AGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGTGATGCAAGCCCTCTGATATG 315
Db 298 AGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGTGATGCAAGCCCTCTGATATG 357
QY 316 AGAAACAAGAGCTTGTCTAAAGATACCTCAAAAAGAGATGATGCAACCAAGATCTGACT 375
Db 358 AGAAACAAGAGCTTGTCTAAAGATACCTCAAAAAGAGATGATGCAACCAAGATCTGACT 417
QY 376 GAGCAGTAGAGGTAGGAGATAAAGATCGGATTCGAAATTTGAGCAAGAGGACTGTAAAG 435
Db 418 GAGCAGTAGAGGTAGGAGATAAAGATCGGATTCGAAATTTGAGCAAGAGGACTGTAAAG 477
QY 436 GTCACAAGGCACACACAGCAGATTTGTAACGGCTATTAAAGACTTATGCGGGTTCCTGTT 495
Db 478 GTCACAAGGCACACACAGCAGATTTGTAACGGCTATTAAAGACTTATGCGGGTTCCTGTT 537
QY 496 GTAGAGGCACCTTCTGAGCAGAGACAGAAATGTGACGCCCTTTGCAATAAACGATAAGGTG 555
Db 538 GTAGAGGCACCTTCTGAGCAGAGACAGAAATGTGACGCCCTTTGCAATAAACGATAAGGTG 597
QY 556 TTCGCTGTGCTTCAGAGATATGGACTCCCTTACTTTTGGGGCTCCAGCGTTCTCTGTT 615
Db 598 TTCGCTGTGCTTCAGAGATATGGACTCCCTTACTTTTGGGGCTCCAGCGTTCTCTGTT 657
QY 616 CATTTAATGGATCCCAAGTTCGAAAGAAATACCTGATGGAATTTGATGTTGCAAGGTT 675
Db 658 CATTTAATGGATCCCAAGTTCGAAAGAAATACCTGATGGAATTTGATGTTGCAAGGTT 717
QY 676 TTGGAGGAGCTTGAACCTCACCATGGACAGTTCATTGATTTGTGCAATCCCTGTGATGAT 735
Db 718 TTGGAGGAGCTTGAACCTCACCATGGACAGTTCATTGATTTGTGCAATCCCTGTGATGAT 777
QY 736 GACTATTGTGATAGCATCAAAAGGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCAA 795
Db 778 GACTATTGTGATAGCATCAAAAGGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCAA 837
QY 796 CATGGTCCATAGAAGCATCTTGGAGAAATCTTAATAAGACAGATATCAAAATTCCTGAG 855
Db 838 CATGGTCCATAGAAGCATCTTGGAGAAATCTTAATAAGACAGATATCAAAATTCCTGAG 897
QY 856 GACTGGCTTACCAAGAGCTCGAGGCTTGTTCAGAGAGCCTAATGTCACATTTGGATAT 915

```
Db 898 GACTGGCCTTACCAAGAGCTCGACGCTTGTTCAGGAGCGCTAATGTCACATTGGATATT 957
QY 916 CCTGAGCTAAATGACCTGCACCTCATGAGGAGGTCTCATAGTTTCCTGGTAAAGAT 975
Db 958 CCTGAGCTAAATGACCTGCACCTCATGAGGAGGTCTCATAGTTTCCTGGTAAAGAT 1017
QY 976 AATGGTTTCAACGAAGATCGGCTGACAAAGCCCATAGAGAAGATCAAAATCTGCCAAGAAT 1035
Db 1018 AATGGTTTCAACGAAGATCGGCTGACAAAGCCCATAGAGAAGATCAAAATCTGCCAAGAAT 1077
QY 1036 AATCGTCCCAAGAGAGCTCGAGTCCTTTTCAAGCCAACTGCCACACATCAGCACCG 1095
Db 1078 AATCGTCCCAAGAGAGCTCGAGTCCTTTTCAAGCCAACTGCCACACATCAGCACCG 1137
QY 1096 CTAAACGGAAGAGAGCTTCGGATTAACAAAGCAAGCAGCTGCGAACAAGAAACAAAG 1155
Db 1138 CTAAACGGAAGAGAGCTTCGGATTAACAAAGCAAGCAGCTGCGAACAAGAAACAAAG 1197
QY 1156 GCTGGTGGAAAGAAATAATCTTGGATGCTTGATGTACAACTACGACTACGAAAGCAG 1215
Db 1198 GCTGGTGGAAAGAAATAATCTTGGATGCTTGATGTACAACTACGACTACGAAAGCAG 1257
QY 1216 CGTGGCGGTGATCACTTCGCTTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTTGGT 1275
Db 1258 CGTGGCGGTGATCACTTCGCTTAGATTATTTAACTCCCTGTTTAACTCAGAGCTTTGGT 1317
QY 1276 AAAAGTT 1282
Db 1318 AAAAGTT 1324

RESULT 4
LOCUS AR152406 1478 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 7 from patent US 6232527.
ACCESSION AR152406
VERSION AR152406.1 GI:15118456
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1478)
AUTHORS Mahajan, P. B.
TITLE Maize Rad2/FEN-1 orthologues and uses thereof
JOURNAL Patent: US 6232527-A 7 15-MAY-2001;
FEATURES
Source
Location/Qualifiers
1..1478
/organism="unknown"
BASE COUNT 463 a 302 c 365 g 348 t
ORIGIN

Query Match 61.5%; Score 850; DB 6; Length 1478;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1200; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 16 GCCACAGCCGCGCAGAGATGGGATCAAGGTTTGACCAAACTCGTGGCGCAAT 75
Db 76 GCCACAGCCGCGCAGAGATGGGATCAAGGTTTGACCAAACTCGTGGCGCAAT 135
QY 76 GCGCCCAAGCGATGAAGAGCAGAAGTTCGAGAGCTACTTCGGCGCGCAAAATCGCGTC 135
Db 136 GCGCCCAAGCGATGAAGAGCAGAAGTTCGAGAGCTACTTCGGCGCGCAAAATCGCGTC 195
QY 136 GAGCCAGCATGAGCATATACCAAGTTCCTGATTGTAGTTGAAGGACAGGCGATGGAACCT 195
Db 196 GAGCCAGCATGAGCATATACCAAGTTCCTGATTGTAGTTGAAGGACAGGCGATGGAACCT 255
QY 196 CTCACAAATCAAGCTGGTGAAGTACTAGTCAATTTGCAAGGAATGTTCAACCGGACAATA 255
Db 256 CTCACAAATCAAGCTGGTGAAGTACTAGTCAATTTGCAAGGAATGTTCAACCGGACAATA 315
QY 256 AGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGTGATGTCAGAGCTTCCTGATG 315
```

```
Db 316 AGATTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTTCATGGCAAGCTCCTCTGATG 375
QY 316 RAGAACAGAGAGCTTGTCTAAAGATCTCAAAAGAGATGATCAACCAAGATCTTGACT 375
Db 376 RAGAACAGAGAGCTTGTCTAAAGATCTCAAAAGAGATGATCAACCAAGATCTTGACT 435
QY 376 GAGGAGCTAGAGTAGGAGATTAAGATGCGATTGAAAAATTTGAGCAAGAGAGCTGTAAAG 435
Db 436 GAGGAGCTAGAGTAGGAGATTAAGATGCGATTGAAAAATTTGAGCAAGAGAGAGCTGTAAAG 495
QY 436 GTCACAAGGCAACCAACGAGAGATTGTAAGCGCTATTAAAGACTTATGSGGGTTCCTGTT 495
Db 496 GTCACAAGGCAACCAACGAGAGATTGTAAGCGCTATTAAAGACTTATGSGGGTTCCTGTT 555
QY 496 GTAGAGGACCTTCTGAAGCAGAAGCAGAAATGTCAGCCCTTTTGCATAAACGATAAGGTG 555
Db 556 GTAGAGGACCTTCTGAAGCAGAAGCAGAAATGTCAGCCCTTTTGCATAAACGATAAGGTG 615
QY 556 TTCGCTGTTGCTTCAGAGATATGGAGCTCCCTTACTTTTGGGGCTCCACGGTTCCTTCGT 615
Db 616 TTCGCTGTTGCTTCAGAGATATGGAGCTCCCTTACTTTTGGGGCTCCACGGTTCCTTCGT 675
QY 616 CATTTAATGATCCCAAGTTCACAGAAAATACCTGTGTATGGAATTTGATGTTGCCAAGTT 675
Db 676 CATTTAATGATCCCAAGTTCACAGAAAATACCTGTGTATGGAATTTGATGTTGCCAAGTT 735
QY 676 TTGGAGGAGCTTGAACCTCACCAGTTCATGATTGATTTGTCATCTCTGTCGATGCT 735
Db 736 TTGGAGGAGCTTGAACCTCACCAGTTCATGATTGATTTGTCATCTCTGTCGATGCT 795
QY 736 GACTATTGTATAGCATCAAAAGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCAA 795
Db 796 GACTATTGTATAGCATCAAAAGTATCGGGGGCAACAGCTCTGAAACTTATTCGTCAA 855
QY 796 CATGGGTCATAGAAAGCATCTTGGAGAATCTTAATAAGACAGATATCAAAATTCCTGAG 855
Db 856 CATGGGTCATAGAAAGCATCTTGGAGAATCTTAATAAGACAGATATCAAAATTCCTGAG 915
QY 856 SACTGGCCTTACCAGAGCTCGACGCTTGTTCAGAGGCTTAATGTCATTTGGATATT 915
Db 916 SACTGGCCTTACCAGAGCTCGACGCTTGTTCAGAGGCTTAATGTCATTTGGATATT 975
QY 916 CTTGAGCTAAAATGACCTGCACCTCATGAGAGGCTCTCATAGTTTTCCTGGTAAAGAT 975
Db 976 CTTGAGCTAAAATGACCTGCACCTCATGAGAGGCTCTCATAGTTTTCCTGGTAAAGAT 1035
QY 976 AATGGTTTCAACGAAGATCGGCTGACAAAGCCCATAGAGAAGATCAAAATCTGCCAAGAAT 1035
Db 1036 AATGGTTTCAACGAAGATCGGCTGACAAAGCCCATAGAGAAGATCAAAATCTGCCAAGAAT 1095
QY 1036 AATCGTCCCAAGAGAGCTCGAGTCCTTTTCAAGCCAACTGCCACACATCAGCACCG 1095
Db 1096 AATCGTCCCAAGAGAGCTCGAGTCCTTTTCAAGCCAACTGCCACACATCAGCACCG 1155
QY 1096 CTAAACGGAAGAGAGCTTCGGATTAACAAAGCAAGCAGCTGCGAACAAGAAACAAAG 1155
Db 1156 CTAAACGGAAGAGAGCTTCGGATTAACAAAGCAAGCAGCTGCGAACAAGAAACAAAG 1215
QY 1156 GCTGGTGAAGAGAAATAATCTTGGATGCTTGAATGTCATGTAAGTACGACTACCAAGAGCAG 1215
Db 1216 GCTGGTGAAGAGAAATAATCTTGGATGCTTGAATGTCATGTAAGTACGACTACCAAGAGCAG 1275
QY 1216 CGGTGGC 1222
Db 1276 CGGTGGC 1282

RESULT 5
LOCUS AB021666 1354 bp mRNA linear PLN 25-MAR-2000
DEFINITION Oryza sativa OsFEN-1 mRNA for FEN-1, complete cds.
ACCESSION AB021666
```

```

VERSION      AB021666.2 GI:4587224
KEYWORDS     FEN-1; endonuclease.
SOURCE       Oryza sativa cDNA to mRNA.
ORGANISM     Oryza sativa
REFERENCE    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Ehrhartoideae; Oryzaceae; Oryza.
TITLE        1 (sites)
JOURNAL      Kimura,S., Ueda,T., Hatanaka,M., Takenouchi,M., Hashimoto,J. and
MEDLINE      Sakaguchi,K.
REFERENCE    Plant Mol. Biol. 42 (3), 415-427 (2000)
AUTHORS      20256470
TITLE        Plant homologue of flap endonuclease-1: molecular cloning,
JOURNAL      characterization, and evidence of expression in meristematic
MEDLINE      tissues
AUTHORS      Kimura,S., Ueda,T., Hatanaka,M., Takenouchi,M., Hashimoto,J. and
              Sakaguchi,K.
TITLE        Submitted (21-DEC-1998) Seisuke Kimura, Science University of
JOURNAL      Tokyo, Dept. of Applied Biological Science: 2641 Yamazaki, Noda,
              Chiba 278-8510, Japan (E-mail:16498703@ed.noda.sut.ac.jp,
              Tel:81-471-24-1501(ex.3419), Fax:81-471-23-9767)
COMMENT      On Apr 17, 1999 this sequence version replaced gi:4062866.
FEATURES     Location/Qualifiers
              1..1354
              /organism="Oryza sativa"
              /db_xref="taxon:4530"
              57..1209
              /gene="OsFEN-1"
              67..1209
              /gene="OsFEN-1"
              /note="endonuclease"
              /codon_start=1
              /product="FEN-1"
              /protein_id="BAA36171.1"
              /db_xref="GI:4587225"
              /translation="MGIKGLTKLLADNAPKMKQKFESYFGRRIAVDASMSITQFLI
              VVGRTMETITNAGEVETSLQGMFNRTIRLLEAGIKPVYFDGKPPDLKKQELAKRY
              SKREDATKELTEAVEGDKDAIEKFSKRTKVYTKQHNECKRLFLMGVPVVEAPCEA
              EAECALCINDMYAVASEDMSLTGAPRFLHMDPSKKIPVMEPEVAKVLEELE
              LTMQDFLDLGLGCDYCDISIKIGIGTALKTIRHGSGTSEILENINKDRYQIPEIDWP
              YQEARRFKPNVTLDIPELKNWAPDEGLVFLKENGENDRVTYKAIKIKFAKK
              SSGRLESFFKPVVSTVPLKRRDKSEKPTKAVANKTKGAGKKK"
              1354
polya_site   /note="17 A nucleotides"
BASE COUNT  415 a 286 c 331 g 322 t
ORIGIN
Query Match 4.2%; Score 58; DB 8; Length 1354;
Best Local Similarity 100.0%; Pred. No. 1e-21;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 TGGCGGACAATCGCCCAAGCGGATGAAGGAGCAGAGTTCGAGAGCTACTTCGGCGC 122
|||||
Db 95 TGGCGGACAATCGCCCAAGCGGATGAAGGAGCAGAGTTCGAGAGCTACTTCGGCGC 152
|||||

RESULT 6
AC104713/c AC104713 115907 bp DNA linear HTG 20-DEC-2001
LOCUS Oryza sativa chromosome 5 clone OJ1362G11, *** SEQUENCING IN
DEFINITION PROGRESS ***, 3 ordered pieces.
AC104713
AC104713.1 GI:17940800
VERSION HTG; HTGS_PHASE2.
KEYWORDS Oryza sativa.
SOURCE Oryza sativa.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 115907)
AUTHORS Chow,T.-Y., Hsiung,Y.-I.C., Chen,C.-S., Chen,H.-H., Wu,H.-P.,
              Liu,S.-M., Chao,Y.-T., Chang,S.-J., Chen,T.-R., Chen,Y.-L.,

```

```

Chow,M.-H.J., Hong,Y.-C., Hsiung,J.-N., Hsu,C.-H., Huang,J.-J.,
Kau,P.-I., Lee,M.-C., Leu,H.-L., Lin,S.-J., Wu,L.-F. and Shaw,J.-F.
Oryza sativa BAC OJ1362G11 genomic sequence
2 (bases 1 to 115907)
Chow,T.-Y. and Hsiung,Y.-I.C.
Direct Submission
Submitted (20-DEC-2001) Institute of Botany, Academia Sinica, 128,
Section 2, Academia Road, Nankang, Taipei 11529, Taiwan
The nucleotide sequence of this BAC clone was generated by
combining Monsanto and ASPGC-Taiwan sequencing data.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 3719: contig of 3719 bp in length
* 3720 3739: gap of unknown length
* 3740 110847: contig of 107108 bp in length
* 110848 110867: gap of unknown length
* 110868 115907: contig of 5040 bp in length.
FEATURES     Location/Qualifiers
              1..115907
              /organism="Oryza sativa"
              /cultivar="Nipponbare"
              /db_xref="taxon:4530"
              /chromosome="5"
              /clone="OJ1362G11"
BASE COUNT  32362 a 24964 c 25806 g 32735 t 40 others
ORIGIN
Query Match 4.2%; Score 58; DB 2; Length 115907;
Best Local Similarity 100.0%; Pred. No. 1.2e-21;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 TGGCGGACAATCGCCCAAGCGGATGAAGGAGCAGAGTTCGAGAGCTACTTCGGCGC 122
|||||
Db 8213 TGGCGGACAATCGCCCAAGCGGATGAAGGAGCAGAGTTCGAGAGCTACTTCGGCGC 8156
|||||

RESULT 7
AP004520/c AP004520 110811 bp DNA linear PLN 14-DEC-2001
LOCUS Lotus japonicus genomic DNA, chromosome 5, clone:LJT05P01, TM0048,
DEFINITION complete sequence.
ACCESSION AP004520
VERSION AP004520.1 GI:17736887
KEYWORDS HTG.
SOURCE Lotus japonicus DNA, clone_lib:LJT library clone:LJT05P01.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;
              Lotus.
REFERENCE 1 (sites)
AUTHORS Sato,S., Kaneko,T., Nakamura,Y., Asamizu,E., Kato,T. and Tabata,S.
TITLE Structural Analysis of a Lotus japonicus Genome. I. Sequence
              Features and Mapping of Fifty-six TAC clones which cover the 5.4 Mb
              Regions of the Genome
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 110811)
TITLE Nakamura,Y.
JOURNAL Direct Submission
REFERENCE Submitted (13-DEC-2001) Yasukazu Nakamura, Kazusa DNA Research
              Institute, Department of Plant Gene Research; 1532-3, Yana,
              Kisarazu, Chiba 292-0812, Japan (E-mail:ynakam@kazusa.or.jp,
              URL:http://www.kazusa.or.jp, Tel:81-438-52-3935,
              Fax:81-438-52-3934)
              Location/Qualifiers
FEATURES

```

```

Source
1..110811
/organism="Lotus japonicus"
/db_xref="taxon:34305"
/chromosome="5"
/clone="Lj105P01"
/clone_lib="Lj1 library"
/note="TAC clone:TM0048"
BASE COUNT 35030 a 19786 c 20743 g 35252 t
ORIGIN

      2.2%: Score 30; DB 8; Length 110811;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 TTGAGAGCTACTTCGGCGCGCAAAATCGCC 132
|||||
Db 78321 TTGAGAGCTACTTCGGCGCGCAAAATCGCC 78292

RESULT 8
AF215846 161494 bp DNA linear HTG 08-JUN-2001
LOCUS Homo sapiens chromosome 8 clone RP11-53M11 map 8q12-q13, WORKING
DEFINITION DRAFT SEQUENCE, 9 unordered pieces.
ACCESSION AF215846
VERSION AF215846.3 GI:14329029
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 161494)
Schlhabel, M.B., Baumgart, C., Blechschmidt, K., Dette, M., Jahn, N.,
Lehmann, R., Menzel, U., Polley, A., Reichwald, K., Schudy, A.,
Siddiqui, R., Taudien, S., Wen, G., Rosenthal, A. and Platzer, M.
Chromosome 8 genomic sequence
Unpublished
2 (bases 1 to 161494)
Polley, A., Wen, G., Baumgart, C., Dette, M., Jahn, N., Schlhabel, M.,
Menzel, U. and Rosenthal, A.
Direct Submission
Submitted (14-DEC-1999) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
On Jun 8, 2001 this sequence version replaced gi:8151782.
-----
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
-----
Project Information
Center project name: H247
Center clone name: RP11-53M11
-----
Summary Statistics
Sequencing vector: MJ3; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 140531 bases at least Q40
Consensus quality: 148839 bases at least Q40
Consensus quality: 154469 bases at least Q20
Quality coverage: 5.33 in Q20 bases; sum-of-contigs
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality 10.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces

```

```

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 3227: contig of 3227 bp in length
* 3228 3327: gap of unknown length
* 3328 66838: contig of 63511 bp in length
* 66838 66939: gap of unknown length
* 66939 99679: contig of 32741 bp in length
* 99679 99779: gap of unknown length
* 99779 115422: contig of 15643 bp in length
* 115422 115523: gap of unknown length
* 115523 127635: contig of 12113 bp in length
* 127635 127715: gap of unknown length
* 127715 137815: contig of 9980 bp in length
* 137815 146218: gap of unknown length
* 146218 146319: gap of unknown length
* 146319 148839: contig of 2521 bp in length
* 148839 148940: gap of unknown length
* 148940 161494: contig of 12555 bp in length.
FEATURES
Location/Qualifiers
1..161494
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q12-q13"
/clone="RP11-53M11"
BASE COUNT 45158 a 31887 c 33658 g 49991 t 800 others
ORIGIN

Query Match 2.0%; Score 27; DB 2; Length 161494;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1355 CTGTTTGTGAAAAAATAAAAAAAAA 1381
|||||
Db 134476 CTGTTTGTGAAAAAATAAAAAAAAA 134502

RESULT 9
AL670227/c 220761 bp DNA linear HTG 30-JAN-2002
LOCUS Mus musculus chromosome 4 clone RP23-317N1, *** SEQUENCING IN
DEFINITION PROGRESS ***, in unordered pieces.
ACCESSION AL670227
VERSION AL670227.2 GI:18477084
KEYWORDS HTG; HTGS_PHASE1.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
Plumb, B
Direct Submission
Submitted (23-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 1, 2002 this sequence version replaced gi:18307364.
-----
Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
Project Information
Center project name: BM317N1
-----
Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator Big Dye; 100% of reads
Consensus quality: 219697 bases at least Q40
Consensus quality: 219887 bases at least Q30

```

source
T. 1203

SRMINEELHASWDQPTGCTIIFRNVELSRVQALFQLTEKLSILAESNERSEARLGG
GGDLPPRRDQDYAAAGGGGTSSGRWQDLSSYQTRQSGRAGYGGRALSFNQ
AGSGGYSRGRGGGYQNSRTQGSALRPHGDVSTRMVSRLRGVRA"

BASE COUNT 367 a 266 c 348 g 414 t

ORIGIN

Query Match 1.9%; Score 26; DB 8; Length 1395;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1356 TCTTTTGAAGAAAAA 1381

Db 1368 TCTTTTGAAGAAAAA 1393

RESULT 12

AC068906/c

LOCUS 192363 bp DNA linear ROD 29-JAN-2002
DEFINITION Mus musculus chromosome 1 clone rp23-240p23 strain C57BL/6J,
complete sequence.

ACCESSION AC068906

VERSION AC068906.12 GI:11181790

KEYWORDS HTG.

SOURCE house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 192363)

RAHAI, R., SHAULL, S., YAO, Z., GROSS, K., JONES, C. and ROE, B.A.

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 192363)

RAHAI, R., SHAULL, S., YAO, Z., GROSS, K., JONES, C. and ROE, B.A.

Direct Submission

Submitted (11-MAY-2000) Department of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

3 (bases 1 to 192363)

RAHAI, R., SHAULL, S., YAO, Z., GROSS, K., JONES, C. and ROE, B.A.

Direct Submission

Submitted (16-NOV-2000) Department of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

4 (bases 1 to 192363)

RAHAI, R., SHAULL, S., YAO, Z., GROSS, K., JONES, C. and ROE, B.A.

Direct Submission

Submitted (29-JAN-2002) Department of Chemistry And Biochemistry,

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,

OK 73019, USA

On Nov 16, 2000 this sequence version replaced gi:11142004.

COMMENT

----- Genome Center

Center: Department Of Chemistry And Biochemistry

The University Of Oklahoma

Center code:UOKNOR

FEATURES

source

1. 192363 Location/Qualifiers

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/chromosome="1"

/clone_lib="RPci mouse BAC library 23"

48581 a 48238 c 48036 g 47508 t

Query Match 1.9%; Score 26; DB 10; Length 192363;

Best Local Similarity 100.0%; Pred. No. 0.0055;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1355 CTGTTTTTGAAGAAAAA 1380

Db 43606 CTGTTTTTGAAGAAAAA 43581

RESULT 13

AC108055

LOCUS

DEFINITION

Homo sapiens chromosome 4 clone RP11-357L18, *** SEQUENCING IN

PROGRESS ***, 12 unordered pieces.

AC108055

AC108055.2 GI:18464305

HTG; HTGS_PHASE1.

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 193026)

Waterston, R.H.

The sequence of Homo sapiens clone

Unpublished

2 (bases 1 to 193026)

Waterston, R.H.

Direct Submission

Submitted (24-JAN-2002) Genome Sequencing Center, Washington

University School of Medicine, 444 Forest Park Parkway, St. Louis,

MO 63108, USA

On Feb 1, 2002 this sequence version replaced gi:18308870.

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: http://genome.wustl.edu/gsc/index.shtml

----- Project Information -----

Center project name: H_NH0357L18

----- Summary Statistics -----

Sequencing vector: p13; 3%

Sequencing vector: plasmid; 97%

Chemistry: Dye-terminator ET; 0% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 189039 bases at least Q40

Consensus quality: 190126 bases at least Q30

Consensus quality: 190624 bases at least Q20

NOTE: This is a 'working draft' sequence. It currently

consists of 12 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 1100: contig of 1100 bp in length

1101 1200: gap of unknown length

1201 3298: contig of 2098 bp in length

3299 3398: gap of unknown length

3399 6107: contig of 2709 bp in length

6108 6207: gap of unknown length

6208 10091: contig of 3884 bp in length

10092 10191: gap of unknown length

10192 14328: contig of 4137 bp in length

14329 14429: gap of unknown length

14429 17963: contig of 3535 bp in length

17964 18063: gap of unknown length

18064 25610: contig of 7547 bp in length

25611 25710: gap of unknown length

25711 42695: contig of 16985 bp in length

42696 42795: gap of unknown length

42796 61794: contig of 18999 bp in length

61795 92392: contig of 30498 bp in length

92393 92492: gap of unknown length

92493 121045: contig of 28553 bp in length

121046 131145: gap of unknown length

131146 193026: contig of 71881 bp in length.

FEATURES	source	Location/Qualifiers	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
misc_feature		1..193026 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="4" /clone="RP11-357L18"	MO 63108, USA	5 (bases 1 to 196367)	Waterston,R. Direct Submission Submitted (09-MAY-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA	On Oct 18, 2000 this sequence version replaced gi:7574982..	
misc_feature		1..1100 /note="assembly_name:Contig26"				Center: Washington University Genome Sequencing Center	
misc_feature		1201..3298 /note="assembly_name:Contig27"				Center code: WUGSC	
misc_feature		3399..6107 /note="assembly_name:Contig28"				Web site: http://genome.wustl.edu/gsc	
misc_feature		6208..10091 /note="assembly_name:Contig29"				Contact: sapiens@watson.wustl.edu	
misc_feature		10192..114328 /note="assembly_name:Contig30"				----- Summary Statistics -----	
misc_feature		14429..17963 /note="assembly_name:Contig31"				Center project name: H_NH0795002	
misc_feature		18064..25610 /note="assembly_name:Contig32"				-----	
misc_feature		25711..42695 /note="assembly_name:Contig33"					
misc_feature		42796..61794 /note="assembly_name:Contig34"					
misc_feature		61895..92392 /note="assembly_name:Contig35"					
misc_feature		92493..121045 /note="assembly_name:Contig36"					
misc_feature		121146..193026 /note="assembly_name:Contig37"					
BASE COUNT	64539 a	31499 c 32054 g 63809 t	1125 others				
ORIGIN							
Query Match		1.9%; Score 26; DB 2: Length 193026;					
Best Local Similarity		100.0%; Pred. No. 0.0055;					
Matches	26; Conservative	0; Mismatches	0; Indels	0; Caps	0;		
Qy	1356	TGTTTTTGAATAAAAAAAAAAAAAA 1381					
Db	110702	TGTTTTTGAATAAAAAAAAAAAAAA 110727					
RESULT 14							
AC020740							
LOCUS	AC020740	196367 bp	DNA	linear	PRI 09-MAY-2001		
DEFINITION	Homo sapiens BAC clone RP11-79502	from 4,	complete sequence.				
ACCESSION	AC020740						
VERSION	AC020740.5	GI:10864253					
KEYWORDS	HTG.						
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	1	(bases 1 to 196367)					
AUTHORS	Sulston,J.E. and Waterston,R.						
TITLE	Toward a complete human genome sequence						
JOURNAL	Genome Res. 8 (11), 1097-1108 (1998)						
MEDLINE	99063792						
REFERENCE	2	(bases 1 to 196367)					
AUTHORS	Sun,H., Stromatt,C. and Barrett,M.						
TITLE	The sequence of Homo sapiens BAC clone RP11-79502						
JOURNAL	Unpublished						
REFERENCE	3	(bases 1 to 196367)					
AUTHORS	Waterston,R.H.						
TITLE	Direct Submission						
JOURNAL	Submitted (08-JAN-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA						
REFERENCE	4	(bases 1 to 196367)					
AUTHORS	Waterston,R.H.						
TITLE	Direct Submission						
JOURNAL	Submitted (18-OCT-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA						

```
repeat_region /rpt_family="Alu"
8236. .8431 /rpt_family="ERV"
repeat_region /rpt_family="ERV"
8634. .8941 /rpt_family="MaLR"
repeat_region /rpt_family="MaLR"
9531. .9749 /rpt_family="L1"
repeat_region /rpt_family="L1"
10002. .10311 /rpt_family="L1"
repeat_region /rpt_family="L1"
10433. .11259 /rpt_family="L1"
repeat_region /rpt_family="L1"
12517. .12818 /rpt_family="Alu"
repeat_region /rpt_family="Alu"
13029. .13105 /rpt_family="L2"
repeat_region /rpt_family="L2"
14568. .14866 /rpt_family="Alu"
repeat_region /rpt_family="Alu"
14868. .15074 /rpt_family="MER1_type"
repeat_region /rpt_family="MER1_type"
17334. .17459 /rpt_family="L2"
repeat_region /rpt_family="L2"
17649. .17961 /rpt_family="Alu"
repeat_region /rpt_family="Alu"
20099. .20459 /rpt_family="MaLR"
repeat_region /rpt_family="MaLR"
20464. .20513 /rpt_family="MaLR"
repeat_region /rpt_family="MaLR"
21880. .21998 /rpt_family="MER2_type"
repeat_region /rpt_family="MER2_type"
22027. .22106 /rpt_family="MER2_type"
repeat_region /rpt_family="MER2_type"
22373. .22554 /rpt_family="L2"
repeat_region /rpt_family="L2"
22783. .23313 /rpt_family="L2"
repeat_region /rpt_family="L2"
23314. .23719 /rpt_family="MaLR"
repeat_region /rpt_family="MaLR"
23720. .23773 /rpt_family="L2"
repeat_region /rpt_family="L2"
24060. .24210 /rpt_family="MIR"
repeat_region /rpt_family="MIR"
24872. .25088 /rpt_family="MIR"
repeat_region /rpt_family="MIR"
25181. .25388 /rpt_family="L1"
repeat_region /rpt_family="L1"
26033. .26082 /rpt_family="ERV"
repeat_region /rpt_family="ERV"
26156. .26377 /rpt_family="L2"
repeat_region /rpt_family="L2"
27565. .28081 /rpt_family="L1"
repeat_region /rpt_family="L1"
28135. .28816 /rpt_family="L1"
repeat_region /rpt_family="L1"
28828. .29119 /rpt_family="Alu"
repeat_region /rpt_family="Alu"
29120. .29663 /rpt_family="L1"
repeat_region /rpt_family="L1"
29667. .30631 /rpt_family="L1"
repeat_region /rpt_family="L1"
30867. .31096 /rpt_family="L1"
repeat_region /rpt_family="L1"
31108. .32254 /rpt_family="L1"
repeat_region /rpt_family="L1"
33647. .33943 /rpt_family="Alu"
repeat_region /rpt_family="Alu"
33981. .34094 /rpt_family="L1"
repeat_region /rpt_family="L1"
34428. .34801 /rpt_family="L1"
repeat_region /rpt_family="L1"
35477. .35737 /rpt_family="Alu"
repeat_region /rpt_family="Alu"
36657. .36972 /rpt_family="Alu"

repeat_region 37029. .37455 /rpt_family="ERV"
repeat_region 38753. .39054 /rpt_family="Alu"
repeat_region 39957. .40262 /rpt_family="L1"
repeat_region 40263. .40961 /rpt_family="L1"
repeat_region 41279. .41373 /rpt_family="L2"
repeat_region 44041. .44091 /rpt_family="L1"
repeat_region 46587. .47674 /rpt_family="L1"
repeat_region 49233. .49536 /rpt_family="Alu"
repeat_region 50029. .50197 /rpt_family="MIR"
repeat_region 50439. .50602 /rpt_family="L1"
repeat_region 50624. .50824 /rpt_family="L1"
repeat_region 50834. .51185 /rpt_family="L1"
repeat_region 51258. .51486 /rpt_family="L1"

Query Match 1.9%; Score 26; DB 9; Length 196367;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1356 TGTGTTTTGAAAAA 1381
Db 117442 TGTGTTTTGAAAAA 117467

RESULT 15
AC087134 197515 bp DNA linear HTG 09-DEC-2000
LOCUS Mus musculus clone RP23-146B24, WORKING DRAFT SEQUENCE, 15
DEFINITION unorderd pieces.
ACCESSION AC087134
VERSION AC087134.1 GI:11610860
KEYWORDS HTG; HTGS-PHASEI; HTGS-DRAFT.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 197515)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Mouse
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 197515)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1792984
Center clone name: RPI-23_146B24
-----
Summary Statistics
Consensus quality: 188097 bases at least Q40
Consensus quality: 190985 bases at least Q30
Consensus quality: 192674 bases at least Q20
Estimated insert size: 213000; agarose-fp estimation
Estimated insert size: 196115; sum-of-contigs estimation
Quality coverage: 7.04 in Q20 bases; agarose-fp estimation
Quality coverage: 7.65 in Q20 bases; sum-of-contigs estimation.
```


O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,W.

Direct Submission
 Submitted (05-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jun 5, 2001 this sequence version replaced gi:13376950.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

 Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www.seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu

 Project Information
 Center project name: L7285
 Center clone name: 124_O_19

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

1 20054: contig of 20054 bp in length
 * 20055 20154: gap of 100 bp
 * 20155 23754: contig of 3600 bp in length
 * 23755 23854: gap of 100 bp
 * 23855 40638: contig of 16784 bp in length
 * 40639 40738: gap of 100 bp
 * 40739 46318: contig of 5580 bp in length
 * 46319 46418: gap of 100 bp
 * 46419 58388: contig of 11970 bp in length
 * 58389 58488: gap of 100 bp
 * 58489 108991: contig of 50503 bp in length
 * 108992 109091: gap of 100 bp
 * 109092 112379: contig of 3288 bp in length
 * 112380 112479: gap of 100 bp
 * 112480 135225: contig of 22746 bp in length
 * 135226 135325: gap of 100 bp
 * 135326 149424: contig of 14099 bp in length
 * 149425 149524: gap of 100 bp
 * 149525 170893: contig of 21369 bp in length
 * 170894 170993: gap of 100 bp
 * 170994 210370: contig of 39377 bp in length
 * 210371 210470: gap of 100 bp
 * 210471 234701: contig of 24231 bp in length.

FEATURES

Location/Qualifiers
 1..234701
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="Rp23-124019"
 /clone_lib="RPI-23 Female Mouse BAC"

BASE COUNT 65397 a 51629 c 51390 g 65053 t 1232 others

ORIGIN

Query Match 1.9%; Score 26; DB 2; Length 234701;
 Best Local Similarity 100.0%; Pred. No. 0.0055;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1356 TCGTTTTCAGAAAAAAGAAAAA 1381

Db 58098 TCGTTTTCAGAAAAAAGAAAAA 58073

RESULT 18

HSA420445 HSA420445 1232 bp mRNA linear PRI 23-NOV-2001
 LOCUS Homo sapiens mRNA full length insert cDNA clone EUROIIMAGE 1645612.
 DEFINITION
 ACCESSION AJ420445
 VERSION AJ420445.1 GI:17066309
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Auffray,C., Ansoirge,W., Ballabio,A., Estivill,X., Gibson,K.,
 Lehrach,H., Poustka,A. and Lundberg,J.
 TITLE The European IMAGE consortium for integrated Molecular analysis of
 human gene transcripts
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1232)
 AUTHORS Persson,A.
 TITLE Direct Submission
 JOURNAL Submitted (02-OCT-2001) Persson A., Center for Molecular
 Biotechnology, KTH, SCFAB, Institute of Biotechnology, Roslagsvagen
 30B, 106 91 Stockholm, SWEDEN
 COMMENT This clone is available royalty-free through IMAGE Consortium
 Distributors. IMPORTANT: This sequence represents the full insert
 of this IMAGE cDNA clone. No attempt has been made to verify
 whether this corresponds to the full-length of the original mRNA
 from which it was derived.

FEATURES

Location/Qualifiers
 1..1232
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="EUROIIMAGE 1645612"
 /clone_lib="Soares_testis_NHT"

polyA_signal 268 a 327 c 326 g 311 t
 BASE COUNT
 ORIGIN

Query Match 1.8%; Score 25; DB 9; Length 1232;
 Best Local Similarity 100.0%; Pred. No. 0.018;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTTCAGAAAAAAGAAAAA 1381

Db 1199 GTTTTTCAGAAAAAAGAAAAA 1223

RESULT 19

BC003123

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

NIH-MGC Project URL: http://mgc.nci.nih.gov

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

BC003123 Homo sapiens, clone IMAGE:3502107, mRNA, partial cds.
 BC003123
 BC003123.1 GI:13111904

human.

ORGANISM Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1808)

Strausberg,R.

Direct Submission

Submitted (13-FEB-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: http://mgc.nci.nih.gov

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabh, Farvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 6 Row: d Column: 8
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source

Location/Qualifiers

1. .1808

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clones="IMAGE:3502107"

/tissue_type="Brain, neuroblastoma"

/clone_lib="NIH_MGC_19"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

<1. .1456

/codon_start=2

/product="Unknown (protein for IMAGE:3502107)"

/protein_id="AAH03123.1"

/db_xref="GI:13111905"

/translation="SGRPRRPAQYPARPRPASTAGSAPFPEGMMGCFALQTVDT

ELTADSVEMCPLOGCRHLLACCTYOLRRPDRPAGPONGKGMVEKPEQVRLGRFLYS

FNNNSIHPVVEVQRDTSAILDMKCHIPVAGHALIGLADASGSIOLRLVESEKSH

VLRLSLALAEQCLALSIDMTGKTGRAGDPLKIISSDSTGQLHLLMNEPRQLQ

KVASWQAHQFEAWIAFNWHEIYVSGDDGLRGWDTRVPGFLFTSKRHTMGVCS

IQSPFHRLIATGSDYDEHILLWDTNRMKPOPLADTPVGGVWRKRWHPFHLLLAAC

MHSGFKILNCKAMERQBATVLTSHIPLDSLVYGADWSLLFRSLQRPSPFNSL

GTADTLKASELPTFCHCRDNDEGHARQSGMKPLTEGMRKNGTWLQVATAATR

DCGVNFEADSAFLLATCSFYDHALLHWEWEGN"

418 a 510 c 520 g 360 t

BASE COUNT

ORIGIN

Query Match 1.8%; Score 25; DB 9; Length 1808;

Best Local Similarity 100.0%; Pred. No. 0.019;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTGTGAAAAAAGAAAAA 1381

|||||

Db 1779 GTTTTGTGAAAAAAGAAAAA 1803

|||||

RESULT 20

BC017335

LOCUS

DEFINITION

Homo sapiens, clone MGC:29782 IMAGE:4642600, mRNA, complete cds.

BC017335

ACCESSION

BC017335.1 GI:16878274

VERSION

MGC.

KEYWORDS

human.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1834)

REFERENCE

Strasbourg, R.

Direct Submission

Submitted (05-NOV-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

AUTHORS

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgaps-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>Contact: nisc-mgc@nigri.nih.gov

Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,

Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,

Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,

Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,

Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,

Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 39 Row: f Column: 23

This clone was selected for full length sequencing because it

passed the following selection criteria: Genomescan gene

prediction.

FEATURES

source

Location/Qualifiers

1. .1834

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clones="MGC:29782 IMAGE:4642600"

/tissue_type="Skin, melanotic melanoma."

/clone_lib="NIH_MGC_20"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

<1. .1485

/codon_start=1

/product="Unknown (protein for MGC:29782)"

/protein_id="AAH17335.1"

/db_xref="GI:16878275"

/translation="MMGCFALQTVDTTELTAADVEMCPLOGCRHLLACCTYOLRRPDR

PAGPONGKGMVEKPEQVRLGRFLYSFNNNSIHPVVEVQRDTSAILDMKCHIPVA

CHALLGLADASGSIOLRLVESEKSHVLEPLSLALEQCLALSIDMTGKTGRAGDQ

PLKIISSDSTGQLHLLMNEPRQLQKVASWQAHQFEAWIAFNWHEIYVSGDDG

LLRGWDTRVPGFLFTSKRHTMGVCSIQSPFHRLIATGSDYDEHILLWDTNRMKPOPL

ADTPVGGVWRKRWHPFHLLLAACMHSGFKILNCKAMERQBATVLTSHIPLDSL

VYGADWSLLFRSLQRPSPFNSLGTADTLKASELPTFCHCRDNDEGHARQ

SGMKPLTEGMRKNGTWLQVATAATRCGVNFEADSAFLLATCSFYDHALLHWEWEGN"

417 a 520 c 531 g 366 t

BASE COUNT

ORIGIN

Query Match 1.8%; Score 25; DB 9; Length 1834;

Best Local Similarity 100.0%; Pred. No. 0.019;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTGTGAAAAAAGAAAAA 1381

|||||

Db 1808 GTTTTGTGAAAAAAGAAAAA 1832

|||||

RESULT 21

HSM801117

LOCUS

DEFINITION

Homo sapiens mRNA; cDNA DKF2p434L108 (from clone DKF2p434L108).

HSM801117

ACCESSION

AL117584

VERSION

AL117584.1 GI:5912144

KEYWORDS

human.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2830)

REFERENCE

Ottewaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and

Wiemann, S.

Direct Submission

Submitted (15-SEP-1999) MIPS, Am Klopferspitz 18a, D-82152

Martinsried, GERMANY

COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp434L108) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES

Location/Qualifiers

1. .2830

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="DKFZp434L108"

/tissue_type="testis"

/clone_lib="434 (synonym: htes3). Vector pSport1; host DH10B; sites NotI + SalI"

/dev_stage="adult"

1. .1694

/gene="DKFZp434L108"

1. .1694

/note="Homo sapiens GLE1, partly"

/codon_start=3

/product="hypothetical protein"

/protein_id="CAB56006.1"

/db_xref="GI:5912145"

/translation="ELVHRMKGTGLRQEOERKVOALSEWSEOLKRFDEKELK QHKEFDLREVMEKSSREALCHQEKIKAERHRAKILNKLREAEQQRVKQAEQLR KEEGILRLIALYAEAELOLSOOLDAEOHALLKLVDLAAFTQNGNQLCSLISGIR ASSSEYPTAEAEAEAEALREWRDLNMGIEITRACDKRRQDEEEAQVKLQAOQ QCPPEAHKEFPAPSGCGKQEDLVKQVODITMOWYQLODASMOCVLTFFGLTNSK DSQAKIKMDLQAAATIPVSQISTIAQSKLKEIFDIHSLSGKPVQSGRGVSVTLN POGDLVOYKLAERFKGEEVASHHEAAFPAAVAVSGIWEHPRVGLDLIAHLHKK CPYSVPVPTFKGMALEDYQRMGLQVQKDSKVEQODNFKRMGMIRLYAAIOLRW PYGNRTHPEHGHGHWRLAQILNMEPLSDVTATLLEFLWCGNALMKQYQVQFWK MLILKEDYFPREAITSSQMGSFIRLQKQFLEKLCLOHQKDIPVPGFLTSEWRS"

polyA_signal 2879. .2794

polyA_site 2807

BASE COUNT 800 a 622 c 724 g 684 t

ORIGIN

Query Match 1.8%; Score 25; DB 9; Length 2830;

Best Local Similarity 100.0%; Pred. No. 0.019;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTGTGAAAAAATAAAAAA 1381

Db 2800 GTTTTGTGAAAAAATAAAAAA 2824

RESULT 22

AC035150/c

LOCUS Homo sapiens chromosome 19, BAC C1978SKB_99E8 (BC67347), complete sequence.

DEFINITION AC035150

VERSION AC035150.1 GI:7459859

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 90141)

AUTHORS Kodyiantri, V., Ge, Y., Severin, J., Krummel, G.K., Gordon, L., Shannon, M., Brower, A., Olsen, A.S. and Smith, L.M.

TITLE Sequence analysis of a 1mb region in 19q13.2 containing a zinc finger gene cluster

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 90141)

AUTHORS Kodyiantri, V., Ge, Y., Severin, J., Krummel, G.K., Gordon, L., Shannon, M., Brower, A., Olsen, A.S. and Smith, L.M.

TITLE Direct Submission

JOURNAL Submitted (06-APR-2000) Department of Chemistry, University of Wisconsin, 1101 University Ave., Madison, WI 53706, USA

COMMENT Map and sequence oriented from q centromere to telomere. BC67347 (C1978SKB_99E8) is adjacent to BC654370 (C19B-H1.2160021) on the left with an estimated gap of less than 1 kb and overlaps F22273 (LLNU-F.180H5, AC018725) on the right from bases 84785 to 90141 of this accession. Additional chr 19 map and sequence information are available at: <http://www.bio.lni.lni.gov/genome/genome.html>.

FEATURES

Location/Qualifiers

1. .90141

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="19"

/map="19q13.2"

/clone="BC67347"

/clone_lib="C1978SKB"

/complement(1. .1852)

/rpt_family="Tigger1"

1863. .3973

repeat_region

1863. .3973

repeat_region

complement(3992. .4267)

repeat_region

/rpt_family="L1MD3"

complement(4270. .4608)

repeat_region

/rpt_family="MLTIE2"

complement(4713. .4845)

repeat_region

/rpt_family="L1M3"

complement(4846. .5144)

repeat_region

/rpt_family="AluY"

complement(5145. .5349)

repeat_region

/rpt_family="L1M3"

complement(5350. .5879)

repeat_region

/rpt_family="L1PA8"

5880. .5916

repeat_region

/rpt_family="L1PA8"

complement(5917. .5965)

repeat_region

/rpt_family="MLTIE"

complement(5952. .6033)

repeat_region

/rpt_family="MLTIE"

complement(6104. .6455)

repeat_region

/rpt_family="L1MC4"

6456. .6672

repeat_region

/rpt_family="L1MC4"

complement(6674. .7077)

repeat_region

/rpt_family="L1ME"

7114. .7226

repeat_region

/rpt_family="GA-rich"

complement(7309. .7590)

repeat_region

/rpt_family="L1R42"

7594. .7999

repeat_region

/rpt_family="L1ME3"

complement(8153. .8439)

repeat_region

/rpt_family="AluJb"

9547. .10368

repeat_region

/rpt_family="L1MB7"

10368. .10692

repeat_region

/rpt_family="AluSg"

complement(11160. .11316)

repeat_region

/rpt_family="MER5A"

12229. .12248

repeat_region

/rpt_family="(T)n"

12347. .12653

misc_feature

/note="BLASTN similarity to AA523472 (24. .330); match 0.99; database searched: est; n167e02.s1 NCI_CGAP_Pr12

0.99; database searched: est; n167e02.s1 NCI_CGAP_Pr12

repeat_region

12859. .12888

repeat_region

/rpt_family="(CA)n"

13451. .13539

repeat_region

/rpt_family="L2"

13939. .14129

repeat_region

/rpt_family="L1MC/D"

14183. .14489

repeat_region

/rpt_family="AluSx"

14501. .15159

repeat_region

sequence is ambiguous, there is an annotation using the 'unsure' feature key.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at
<http://www.sanger.ac.uk/projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr6>
 RP4-662A9 is from the library RPCI-4 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>
 VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone RP4-662A9 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
 The true right end of clone RP4-662A9 is at 109088 in this sequence. The true left end of clone RP11-373A10 is at 111 in this sequence. The true right end of clone RP3-323P13 is at 100 in this sequence.

FEATURES

Location/Qualifiers

```

source
1..109088
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="6"
  /clone="RP4-662A9"
  /clone_lib="RPCI-4"
  complement(1..117)
  /note="match: GSS: Em:AQ053756"
  complement(1..100)
  /note="match: GSS: Em:AQ469668"
  complement(1..70)
  /note="match: STS: Em:HS323P13T"
  140..623
  /note="match: GSS: Em:AQ529745"
  144..499
  /note="match: GSS: Em:AQ530208"
  complement(154..552)
  /note="match: GSS: Em:AQ027355"
  complement(187..496)
  /note="match: GSS: Em:AQ035558"
  complement(376..771)
  /note="match: STS: Em:HS196E19T"
  complement(1316..1782)
  /note="match: GSS: Em:AQ312880"
  complement(1372..1787)
  /note="match: GSS: Em:AQ197093"
  5398..5429
  /note="16 copies 2 mer aa 87% conserved"
  5813..6050
  /note="MLRIG repeat: matches 20..266 of consensus"
  6078..6356
  /note="MLRIG repeat: matches 270..541 of consensus"
  6532..6698
  /note="MLR1-INTERNAL repeat: matches 292..468 of consensus"
  complement(9447..9891)
  /note="match: GSS: Em:AQ340514"
  13728..13783
  /note="8 copies 7 mer aaagaag 80% conserved"
  14023..14521
  /note="L1 repeat: matches 2915..3440 of consensus"
  L1 repeat: matches 2915..3440 of consensus"
  14819..14880
  /note="L1 repeat: matches 2857..2915 of consensus"
  L1 repeat: matches 2857..2915 of consensus"
  14890..15185
  /note="MER2 repeat: matches 6..345 of consensus"
  MER2 repeat: matches 6..345 of consensus"
  15187..15410
  /note="L1 repeat: matches 2630..2853 of consensus"

```

```

misc_feature
16460..16753
  /note="match: GSS: Em:AQ336851"
repeat_region
16642..16723
  /note="LRL16C repeat: matches 301..385 of consensus"
repeat_region
17330..17601
  /note="136 copies 2 mer tt 54% conserved"
misc_feature
18019..18453
  /note="match: GSS: Em:AQ031513"
repeat_region
20213..20347
  /note="MER45B repeat: matches 1..173 of consensus"
misc_feature
21331..21939
  /note="match: GSS: Em:AQ001839"
misc_feature
21355..21865
  /note="match: GSS: Em:AQ005578"
repeat_region
22254..22624
  /note="THE1C repeat: matches 1..371 of consensus"
THE1C repeat: matches 1..371 of consensus"
22625..23849
  /note="THE1C-internal repeat: matches 322..1580 of consensus"
repeat_region
23842..23986
  /note="THE1C-internal repeat: matches 1..144 of consensus"
THE1C-internal repeat: matches 1..144 of consensus"
23987..24354
  /note="THE1C repeat: matches 1..371 of consensus"
THE1C repeat: matches 1..371 of consensus"
25515..25558
  /note="11 copies 4 mer tgtg 90% conserved"
25616..26117
  /note="L1M4 repeat: matches 5791..6300 of consensus"
L1M4 repeat: matches 5791..6300 of consensus"
26144..26570
  /note="L1M4 repeat: matches 5290..5721 of consensus"
L1M4 repeat: matches 5290..5721 of consensus"
26577..27342
  /note="L1P13 repeat: matches 5355..6147 of consensus"
L1P13 repeat: matches 5355..6147 of consensus"
27648..30820
  /note="L1P13 repeat: matches 5393..6147 of consensus"
27648..27946
  /note="L1P13 repeat: matches 2299..5393 of consensus"
27946..30820
  /note="L1P13 repeat: matches 5062..5355 of consensus"
31118..32433
  /note="L1P13 repeat: matches 2313..5161 of consensus"
L1P13 repeat: matches 2313..5161 of consensus"
32480..34003
  /note="L1P12 repeat: matches 1002..2299 of consensus"
L1P12 repeat: matches 1002..2299 of consensus"
34306..35200
  /note="L1 repeat: matches 1417..214 of consensus"
L1 repeat: matches 1417..214 of consensus"
35198..36378
  /note="L1M4 repeat: matches 4922..6184 of consensus"
complement(35202..35601)
  /note="match: GSS: Em:B70379"
35289..36378
  /note="L1M4 repeat: matches 4922..6018 of consensus"
36372..36768
  /note="L1 repeat: matches 3985..4440 of consensus"
L1 repeat: matches 3985..4440 of consensus"
36775..36843
  /note="L1 repeat: matches 3867..3935 of consensus"
L1 repeat: matches 3867..3935 of consensus"
40940..41308
  /note="THE1B repeat: matches 1..364 of consensus"
THE1B repeat: matches 1..364 of consensus"
41094..41666
  /note="match: STS: Em:G63667"
43211..43846
  /note="MER77 repeat: matches 18..636 of consensus"

```



```

* * *
63790 63889: gap of unknown length
63890 66981: contig of 3092 bp in length
66981 67081: gap of unknown length
67081 70048: contig of 2967 bp in length
70048 70148: gap of unknown length
70148 74058: contig of 3910 bp in length
74058 74158: gap of unknown length
74158 76571: contig of 2413 bp in length
76571 76671: gap of unknown length
76671 81964: contig of 5293 bp in length
81964 82064: gap of unknown length
82064 85537: contig of 3473 bp in length
85537 85637: gap of unknown length
85637 89225: contig of 3288 bp in length
89225 89025: gap of unknown length
89025 92357: contig of 3332 bp in length
92357 92457: gap of unknown length
92457 95715: contig of 3258 bp in length
95715 95815: gap of unknown length
95815 99089: contig of 3274 bp in length
99089 101633: contig of 2444 bp in length
101633 101733: gap of unknown length
101733 105071: contig of 3338 bp in length
105071 105171: gap of unknown length
105171 107293: contig of 2122 bp in length
107293 107393: gap of unknown length
107393 109288: contig of 1895 bp in length
109288 109388: gap of unknown length
109388 111367: contig of 1979 bp in length
111367 111467: gap of unknown length
111467 113189: contig of 1722 bp in length
113189 113289: gap of unknown length
113289 115104: contig of 1815 bp in length
115104 115204: gap of unknown length
115204 1172507: contig of 2053 bp in length
1172507 117357: gap of unknown length
117357 119053: contig of 1596 bp in length
119053 119053: gap of unknown length
119053 120865: contig of 1812 bp in length
120865 120965: gap of unknown length
120965 121997: contig of 1032 bp in length
121997 122097: gap of unknown length
122097 123797: contig of 1699 bp in length
123797 123897: gap of unknown length
123897 126007: contig of 2111 bp in length
126007 128300: contig of 2193 bp in length
128300 128401: gap of unknown length
128401 130016: contig of 1616 bp in length
130016 130116: gap of unknown length
130116 132181: contig of 2065 bp in length
132181 132281: gap of unknown length
132281 133924: contig of 1643 bp in length
133924 134025: gap of unknown length
134025 135827: contig of 1803 bp in length
135827 135927: gap of unknown length
135927 137312: contig of 1385 bp in length
137312 137412: gap of unknown length
137412 138911: contig of 1499 bp in length
138911 139011: gap of unknown length
139011 141070: contig of 2059 bp in length
141070 141170: gap of unknown length
141170 142389: contig of 1219 bp in length
142389 142490: gap of unknown length
142490 143597: contig of 1108 bp in length
143597 143697: gap of unknown length
143697 145315: contig of 1618 bp in length
145315 145415: gap of unknown length
145415 146958: contig of 1543 bp in length
146958 147059: gap of unknown length
147059 148210: contig of 1152 bp in length.
* * *
Location/Qualifiers

Source
1. .148210
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-121A10"
BASE COUNT 35495 a 36628 c 37058 g 34422 t 4607 others
ORIGIN

Query Match 1.8%; Score 25; DB 2; Length 148210;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTGTGAAAAA 1381
|||||
Db 138750 GTTTTGTGAAAAA 138774
|||||

RESULT 25
AC024144/c 160434 bp DNA linear HTG 25-JAN-2002
LOCUS Mus musculus chromosome 2 clone RP23-106A3 strain C57BL6/J, WORKING
DEFINITION DRAFT SEQUENCE, 174 unordered pieces.
ACCESSION AC024144 GI:18092972
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Fukuoka; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 160434)
Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,
Gordon, M., Goltz, J.S. and Kucherlapati, R.
High Throughput Mouse Sequencing
Unpublished
2 (bases 1 to 160434)
Montgomery, K.T., Grills, G., Han, J., Lee, E., Long, J., Pomerantz, R.,
Ioshikhes, I.P., Shim, C., Decker, J., Thomas, E., Perera, A.,
Gordon, M., Goltz, J.S. and Kucherlapati, R.
Direct Submission
Submitted (25-FEB-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
Bronx, NY 10461, USA
On Jan 9, 2002 this sequence version replaced gi:11094421.
-----Genome Center
Center: Harvard Partners Genome Center
Web site:
http://wchanning.bwh.harvard.edu:9088/hpcgg/jsp/hpcgg/Sequence/mous
e.html
Contact: gntmecapecod.bwh.harvard.edu
-----Summary Statistics
Center project name: ABR
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 100%
*Consensus quality: 34047 at least Q20
*Consensus quality: 30434 at least Q30
*Consensus quality: 24749 at least Q40
Estimated insert size: agarose-FP - N/A
**Estimated insert size: 156974 - sum-of-contigs
Quality coverage: agarose-FP - N/A
Quality coverage: 0.3 x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 174 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 871: contig of 871 bp in length
* 872 892
* 891: gap of unknown length
* 1463: contig of 572 bp in length

```


* 1464 1483: gap of unknown length
* 1484 2078: contig of 595 bp in length
* 2079 2098: gap of unknown length
* 2099 3015: contig of 917 bp in length
* 3016 3035: gap of unknown length
* 3036 3873: contig of 837 bp in length
* 3873 3893: gap of unknown length
* 3893 4674: contig of 782 bp in length
* 4675 5529: gap of unknown length
* 4695 5529: contig of 835 bp in length
* 5530 5550: gap of unknown length
* 5550 6364: contig of 815 bp in length
* 6365 6384: gap of unknown length
* 6385 7153: contig of 775 bp in length
* 7160 7179: gap of unknown length
* 7180 8306: contig of 1127 bp in length
* 8307 8326: gap of unknown length
* 8327 9018: contig of 692 bp in length
* 9019 9038: gap of unknown length
* 9039 9877: contig of 839 bp in length
* 9878 9898: gap of unknown length
* 9898 10581: contig of 684 bp in length
* 10582 10601: gap of unknown length
* 10602 11319: contig of 718 bp in length
* 11320 11339: gap of unknown length
* 11340 12085: contig of 746 bp in length
* 12086 12105: gap of unknown length
* 12106 12924: contig of 819 bp in length
* 12925 12944: gap of unknown length
* 12945 14043: contig of 1101 bp in length
* 14046 14063: gap of unknown length
* 14066 14782: contig of 717 bp in length
* 14783 14803: gap of unknown length
* 14803 15545: contig of 743 bp in length
* 15546 15565: gap of unknown length
* 15566 15663: contig of 1098 bp in length
* 15664 16688: gap of unknown length
* 16688 17551: contig of 868 bp in length
* 17552 17571: gap of unknown length
* 17572 18463: contig of 892 bp in length
* 18464 18483: gap of unknown length
* 18484 19273: contig of 790 bp in length
* 19274 19293: gap of unknown length
* 19294 20193: contig of 900 bp in length
* 20194 20213: gap of unknown length
* 20214 21145: contig of 932 bp in length
* 21146 21165: gap of unknown length
* 21166 22003: contig of 838 bp in length
* 22004 22023: gap of unknown length
* 22024 22900: contig of 877 bp in length
* 22901 22920: gap of unknown length
* 22921 23791: contig of 871 bp in length
* 23792 23811: gap of unknown length
* 23812 24710: contig of 899 bp in length
* 24711 24730: gap of unknown length
* 24731 25511: contig of 781 bp in length
* 25512 25531: gap of unknown length
* 25532 26220: contig of 689 bp in length
* 26221 26240: gap of unknown length
* 26241 26894: contig of 654 bp in length
* 26895 26914: gap of unknown length
* 26915 27828: contig of 914 bp in length
* 27829 27848: gap of unknown length
* 27849 28751: contig of 903 bp in length
* 28752 28771: gap of unknown length
* 28772 29504: contig of 733 bp in length
* 29505 29524: gap of unknown length
* 29525 30444: contig of 920 bp in length
* 30445 30464: gap of unknown length
* 30465 31394: contig of 930 bp in length
* 31395 31414: gap of unknown length
* 31415 32319: contig of 905 bp in length
* 32320 32333: gap of unknown length

32340 33169: contig of 830 bp in length
33170 33189: gap of unknown length
33190 34021: contig of 832 bp in length
34022 34041: gap of unknown length
34042 34901: contig of 860 bp in length
34902 34921: gap of unknown length
34922 35732: contig of 811 bp in length
35733 35752: gap of unknown length
35753 36850: contig of 1098 bp in length
36851 36870: gap of unknown length
36871 37853: contig of 983 bp in length
37854 37873: gap of unknown length
37874 38881: contig of 1008 bp in length
38882 38901: gap of unknown length
38902 39800: contig of 899 bp in length
39801 39820: gap of unknown length
39821 40559: contig of 739 bp in length
40560 40579: gap of unknown length
40580 41300: contig of 721 bp in length
41301 41320: gap of unknown length
41321 42139: contig of 819 bp in length
42140 42159: gap of unknown length
42160 42851: contig of 692 bp in length
42852 42871: gap of unknown length
42852 43550: contig of 679 bp in length
43551 43570: gap of unknown length
43571 44057: contig of 487 bp in length
44058 44077: gap of unknown length
44078 44699: contig of 622 bp in length
44700 44719: gap of unknown length
44720 45413: contig of 694 bp in length
45414 45433: gap of unknown length
45434 46062: contig of 629 bp in length
46063 46082: gap of unknown length
46083 47025: contig of 943 bp in length
47026 47045: gap of unknown length
47046 47928: contig of 883 bp in length
47929 47948: gap of unknown length
47949 48683: contig of 735 bp in length
48684 48703: gap of unknown length
48704 49516: contig of 813 bp in length
49517 49536: gap of unknown length
49537 50285: contig of 749 bp in length
50286 50305: gap of unknown length
50306 51121: contig of 816 bp in length
51122 51141: gap of unknown length
51142 51855: contig of 714 bp in length
51856 51875: gap of unknown length
51876 52702: contig of 827 bp in length
52703 52722: gap of unknown length
52723 53387: contig of 665 bp in length
53407 53407: gap of unknown length
53408 54126: contig of 719 bp in length
54127 54146: gap of unknown length
54147 55010: contig of 864 bp in length
55011 55030: gap of unknown length
55031 55716: contig of 686 bp in length
55717 55736: gap of unknown length
55737 56694: contig of 958 bp in length
56695 56714: gap of unknown length
56715 57342: contig of 628 bp in length
57343 57362: gap of unknown length
57363 58223: contig of 861 bp in length
58224 58243: gap of unknown length
58244 59301: contig of 1058 bp in length
59302 59321: gap of unknown length
59322 60251: contig of 930 bp in length
60252 60271: gap of unknown length
60272 61133: contig of 861 bp in length
61133 61152: gap of unknown length
61153 61824: contig of 672 bp in length
61825 61844: gap of unknown length

```

Query Match      1.8%; Score 25; DB 2; Length 160434;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTTCGAAAAAAGAAAAA 1381
|||||
Db 50599 GTTTTTCGAAAAAAGAAAAA 50575
|||||

RESULT 26
AP001884/c
LOCUS      160492 bp      DNA      linear      PRI 06-JUL-2001
DEFINITION Homo sapiens genomic DNA, chromosome 11q clone:RP11-148F13,
             complete sequence.
ACCESSION  AP001884
VERSION     AP001884.4 GI:14625387
KEYWORDS   HTG.
SOURCE     Homo sapiens DNA, clone:RP11-148F13.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 160492)
AUTHORS    Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
JOURNAL    Homo sapiens genomic DNA
REFERENCE  2 (bases 1 to 160492)
AUTHORS    Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
            Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL    Submitted (24-APR-2000) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT    On Jul 6, 2001 this sequence version replaced gi:9844973.
FEATURES   Location/qualifiers
            source          1..160492
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /chromosome="11"
                        /map="11q"
                        /clone="RP11-148F13"
BASE COUNT  48270 a 29600 c 30456 g 52166 t
ORIGIN

Query Match      1.8%; Score 25; DB 9; Length 160492;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTTCGAAAAAAGAAAAA 1381
|||||
Db 33825 GTTTTTCGAAAAAAGAAAAA 33801
|||||

RESULT 27
AC021070/c
LOCUS      162978 bp      DNA      linear      PRI 08-FEB-2001
DEFINITION Homo sapiens 12 BAC RP11-460N10 (Roswell Park Cancer Institute
             Human BAC Library) complete sequence.
ACCESSION  AC021070
VERSION     AC021070.24 GI:11496330
KEYWORDS   HTG.
SOURCE     Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 162978)
AUTHORS    Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-ouman,F.R., Allen,C.,
            Alsbrooks,S.B., Amarantunge,H.C., Are,J.R., Banks,T., Barbara,J.,
            Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
            Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,C., Edwards,C.C.,
            Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,
            Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
            Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,
            Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,
            Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O.,
            Hodgson,A., Hoques,M., Holloway,C., Hollins,B., Homsí,F.,
            Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,
            Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
            Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
            Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Louisged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R.,
            Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P.,
            Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,
            Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,
            Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Roife,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C.,
            Shoohtari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A.,
            Stanke,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A.,
            Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,
            Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,
            Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,
            Vinsanton,C., Watlington,S., Williams,G., Williamson,A.,
            Wlecyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Zhou,J.,
            Zorrilla,S., Kuchelapati,R. and Gibbs,R.
            Direct Submission
            Unpublished
            2 (bases 1 to 162978)
            Worley,K.C.
            Direct Submission
            Submitted (14-JAN-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            3 (bases 1 to 162978)
            Direct Submission
            Submitted (01-DEC-2000) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            4 (bases 1 to 162978)
            Direct Submission
            Submitted (08-FEB-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
            On Dec 1, 2000 this sequence version replaced gi:11465032.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-helpebcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.

```

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

```

----- Summary Statistics -----
Contig length: 162978
Phrap values in estimate: 159321
Average error rate (BCM-Phrap estimate): 6.83397e-05
Fraction of Phrap values less than 40 : 0.0302032
Number of consensus changing edits: 89
Number of N's in consensus : 0

```

```

----- Consensus changing edits -----
Position Original+Context Edited+Context
4472 agaacatcgc(n)caacagagat agaacatcgc(c)caacagagat
4577 attgcagata(n)agtcagagatc attgcagata(a)agtcagagatc
5643 agccagaaa(n)ccctgccttt agccagaaa(a)ccctgccttt
8349 gccacgttag(n)acaatagaaa gccacgttag(t)acaatagaaa
10534 ggttaacggg(n)agtagttaat ggttaacggg(a)agtagttaat
10548 taataacatt(n)ggttttcaca taataacatt(t)ggttttcaca
17057 ggcattttc(n)nnatattgat ggcattttc(c)nnatattgat
17058 gccattttc(n)natattgat gccattttc(c)natattgat
17059 ccattttcn(n)atatgtatc ccattttcn(n)atatgtatc
18529 ttctagttct(n)ttaatgttga ttctagttct(t)ttaatgttga
23183 caccactat(n)ggnagttaan caccactat(t)ggnagttaan
23192 cactcatng(n)agtaaannc cactcatng(c)agtaaannc
23199 ngnagtaaa(n)cnctgcctta ngnagtaaa(a)cnctgcctta
23201 gcnagtaaac(n)ctgccttaga gcnagtaaac(t)ctgccttaga
26044 ggaagccag(n)ggctgtctc ggaagccag(t)ggctgtctc
26196 gaagttatt(n)tgatgataa gaagttatt(t)tgatgataa
37819 ggtgataccc(n)ngcaaacggt ggtgataccc(a)ngcaaacggt
37820 gtgataccn(n)gcaaacggtg gtgataccn(a)gcaaacggtg
37867 agcagacctg(n)ngcnaggag agcagacctg(c)ngcnaggag
37868 gcagacctgn(n)gcnaggagc gcagacctg(c)ngcnaggagc
37871 gaactgngc(n)ngagagcctg gaactgngc(a)ngagagcctg
37872 acctgngcn(n)agagcctga acctgngcn(a)agagcctga
37886 agcctgactg(n)tagaagaaa agcctgactg(t)tagaagaaa
46134 aaatacaaaa(n)nnnaaaaa aaatacaaaa(t)agcttaaaaa
46135 atacaaaaa(n)nnnaaaaaa atacaaaaa(a)agcttaaaaa
46136 tacaaaaan(n)nnnaaaaaa tacaaaaan(g)agcttaaaaa
46137 caaaaaann(n)naaaaaann caaaaaann(t)agcttaaaaa
46138 caaaaaann(n)naaaaaann caaaaaann(c)tagaagaaatg
46139 caaaaaann(n)naaaaaann caaaaaann(t)tagaagaaatg
46142 nnnnaaaa(n)naaaaaann nnnnaaaa(g)aaatagctaa
46147 nnaaaaaa(n)naaaaaann nnaaaaaa(t)gtaagtgaat
46148 naaaaaaa(n)naaaaaann naaaaaaa(g)taagtgaat
46149 naaaaaaa(n)naaaaaann naaaaaaa(t)taagtgaat
46152 naaaaaaa(n)naaaaaann naaaaaaa(t)taagtgaat
46153 aaannnaa(n)nnnttttnt aaannnaa(t)gagatttata
46154 aaannnaa(n)nnntttnt aaannnaa(t)gagatttata
46155 aaannnaa(n)nnntttnt aaannnaa(t)gagatttata
46156 annnaaaaa(n)tttntttt annnaaaaa(a)tttntttt

```

```

46160 aannnnnttt(n)nnnttnaaat aannnnnttt(a)nnnttnaaat
46161 aannnnnttn(n)nnnttnaaat aannnnnttn(t)nnnttnaaat
46162 nnnnnnttn(n)nnnttnaaatg nnnnnnttn(a)nnnttnaaatg
46166 ntntnnntt(n)aaattgcctnn ntntnnntt(a)aaattgcctnn
46175 naaattgcc(n)nttntntttna naaattgcc(t)nttntntttna
46176 naaattgcc(n)nttntntttna naaattgcc(t)nttntntttna
46180 tgccnnntt(n)nttnnaanna tgccnnntt(a)nttnnaanna
46181 gcnntnttn(n)naannaagng gcnntnttn(a)naannaagng
46184 nntntnttn(n)naannaagng nntntnttn(a)naannaagng
46188 tntnttnaa(n)naagngtgt tntnttnaa(t)naagngtgt
46189 tntnttnaa(n)naagngtgt tntnttnaa(t)naagngtgt
46194 nnaannaag(n)ngtctgtagc nnaannaag(t)ngtctgtagc
46985 ttatttttc(n)nnntacaag ttatttttc(a)nnntacaag
46986 ttttttcn(n)nttacaaga ttttttcn(a)nttacaaga
46987 attttccn(n)nttacaaga attttccn(a)nttacaaga
46988 tttttccn(n)nttacaaga tttttccn(a)nttacaaga
48509 aaagaactg(n)nnnttgatt aaagaactg(t)nnnttgatt
48510 aagaactgn(n)nttgattg aagaactgn(a)nttgattg
48511 agaactggnn(n)tttgatttg agaactggnn(t)tttgatttg
48523 ctcatatta(n)ttatgttgt ctcatatta(t)ttatgttgt
48802 agataaatg(n)tcctcatct agataaatg(t)tcctcatct
60518 aacaatagac(n)ctgggactc aacaatagac(t)ctgggactc
64812 ttttttct(n)tttttaata ttttttct(a)tttttaata
78276 agataatg(n)aaaagtgtc agataatg(t)aaaagtgtc
80348 cctcatcaa(n)ataatagctt cctcatcaa(a)ataatagctt
83194 ttttgctc(n)ttgggaaaag ttttgctc(a)ttgggaaaag
83210 aaaagtttt(n)ccagctgata aaaagtttt(a)ccagctgata
83569 cattcattt(n)aaagcccttc cattcattt(a)aaagcccttc
88284 acatatcgaa(t)cttatagat acatatcgaa(a)cttatagat
95491 cagaagaacca(n)tttcaagct cagaagaacca(g)tttcaagct
96116 aggtgagta(n)ctgtgttagc aggtgagta(a)ctgtgttagc
110134 ctctacict(n)ccatbaagc ctctacict(t)ccatbaagc
117584 gatcagatg(n)tgcatatg gatcagatg(t)tgcatatg
117650 tctgttttg(n)accaatacca tctgttttg(a)accaatacca
117701 tadtttgaag(n)caggtagat tadtttgaag(t)caggtagat

Query Match 1.8%; Score 25; DB 9; Length 162978;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTGTGAAAAAATAAAAAA 1381
Db 93696 GTTTTGTGAAAAAATAAAAAA 93672

RESULT 28
AC018421
LOCUS Homo sapiens clone RP11-19F10, WORKING DRAFT SEQUENCE, 10 unordered
pieces.
AC018421
AC018421.3 GI:7248979
VERSION HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 164382)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-19F10
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 164382)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,E.,
Boguslavsky,L., Boukhgaiter,B., Brown,A., Castle,A., Colangelo,M.,
Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K.,
Domino,M., Doyle,M., Fenestor,J., Ferreira,P., FitzHugh,W.,
Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B.,
Heatford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L.,
Karatas,A., Klein,J., Landers,T., Lehoczy,J., Lieu,C., Locke,K.,

```

Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
 Meldrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T.,
 O'Donnell, P., Peterson, K., Pierre, N., Pollara, V., Riley, R.,
 Rothman, D., Roy, A., Santos, R., Severy, P., Stange-Thomann, N.,
 Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
 Zimmer, A. and Zody, M.

TITLE JOURNAL

Submitted (10-DEC-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 15, 2000 this sequence version replaced gi:6649387.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L3727

Center clone name: 19_F_10

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 158002 bases at least Q40

Consensus quality: 160560 bases at least Q30

Consensus quality: 161568 bases at least Q20

Insert size: 170000; agarose-fp

Quality coverage: 5.0 in Q20 bases; agarose-fp

Quality coverage: 5.2 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 10 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 4066: contig of 4066 bp in length

* 4067 4166: gap of 100 bp

* 4167 10688: contig of 6522 bp in length

* 10689 10788: gap of 100 bp

* 10789 17722: contig of 6934 bp in length

* 17723 17822: gap of 100 bp

* 17823 24213: contig of 6391 bp in length

* 24214 24313: gap of 100 bp

* 24314 33775: contig of 9462 bp in length

* 33776 33875: gap of 100 bp

* 33876 47138: contig of 13263 bp in length

* 47139 47238: gap of 100 bp

* 47239 63019: contig of 15781 bp in length

* 63020 63119: gap of 100 bp

* 63120 84855: contig of 21736 bp in length

* 84856 84955: gap of 100 bp

* 84956 116335: contig of 31380 bp in length

* 116336 116435: gap of 100 bp

* 116436 164382: contig of 47947 bp in length.

* Location/Qualifiers

* 1..164382

* /organism="Homo sapiens"

* /db_xref="taxon:9606"

* /clone="RP11-19F10"

* /clone_lib="RCI-11 Human Male BAC"

* 1..4066

* /note="assembly_fragment"

* 4167..10688

* /note="assembly_fragment"

* clone_end:SP6

* vector_side:right"

misc_feature 10789..17722

/note="assembly_fragment"

misc_feature 17823..24213

/note="assembly_fragment"

misc_feature 24314..33775

/note="assembly_fragment"

misc_feature 33876..47138

/note="assembly_fragment"

misc_feature 47239..63019

/note="assembly_fragment"

vector_side:right

63120..84855

/note="assembly_fragment"

misc_feature 84956..116335

/note="assembly_fragment"

misc_feature 116436..164382

/note="assembly_fragment"

BASE COUNT 48477 a 32716 c 31773 g 50512 t 904 others

ORIGIN

Query Match 1.8%; Score 25; DB 2; Length 164382;

Best Local Similarity 100.0%; Pred. No. 0.021;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTTCAGAAAAAATAAAAAA 1381

|||||TTTTTTTTTTTTTTTTTTTT

Db 72640 GTTTTTCAGAAAAAATAAAAAA 72664

RESULT 29

AL589693

LOCUS AL589693 165203 bp DNA linear PRI 06-APR-2001

DEFINITION Human DNA sequence from clone RP11-19F10 on chromosome 6, complete

sequence.

ACCESSION AL589693

VERSION AL589693.3 GI:13277175

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 165203)

AUTHORS Blakey, S.

TITLE Direct Submission

JOURNAL Submitted (06-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Mar 12, 2001 this sequence version replaced gi:13274924.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em: EMBL; Sw:

SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP

database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping

Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr6

RP11-19F10 is from the library RP11-11.1 constructed by the group

of Pieter de Jong. For further details see

FEATURES source

misc_feature

misc_feature


```

* 163634 163733: gap of 100 bp
* 163734 164967: contig of 1234 bp in length
* 164968 165067: gap of 100 bp
* 165068 166611: contig of 1544 bp in length
* 166612 166711: gap of 100 bp
* 166712 167884: contig of 1173 bp in length
* 167885 167984: gap of 100 bp
* 167985 169235: contig of 1251 bp in length
* 169236 169335: gap of 100 bp
* 169336 170829: contig of 1494 bp in length
* 170830 170929: gap of 100 bp
* 170930 172319: contig of 1390 bp in length
* 172320 172419: gap of 100 bp
* 172420 173562: contig of 1143 bp in length
* 173563 173662: gap of 100 bp
* 173663 174977: contig of 1315 bp in length
* 174978 175077: gap of 100 bp
* 175078 176150: contig of 1073 bp in length.

FEATURES
Source
    1..176150
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="18"
        /map="18q11.2"
        /clone="RP11-874A11"
    misc_feature
        1..12769
            /note="assembly_fragment"
    misc_feature
        12870..126395
            /note="assembly_fragment clone_end:Sp6 vector_side:right"

Query Match      1.8%; Score 25; DB 2; Length 176150;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTTCAGAAAAAAGAAAAA 1381
      |||||||
Db 102664 GTTTTTCAGAAAAAAGAAAAA 102640

RESULT 32
AP002407/c
LOCUS      AP002407      176292 bp      DNA      linear      PRI 21-JUN-2001
DEFINITION Homo sapiens genomic DNA, chromosome 11q, clone:RP11-767M8,
             complete sequence.
ACCESSION  AP002407
VERSION     AP002407.3  GI:14717411
KEYWORDS    HTG.
SOURCE      Homo sapiens DNA, clone:RP11-767M8.
ORGANISM    Homo sapiens
            Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
            Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo:
            1 (sites)
            Hattori.M., Ishii.K., Toyoda.A., Taylor.T.D., Hong-Seog.P.,
            Fujiyama.A., Yada.T., Totoki.Y., Watanabe.H. and Sakaki.Y.
            Homo sapiens genomic DNA
            Published Only in DataBase (2000) In press
            2 (bases 1 to 176292)
            Hattori.M., Ishii.K., Toyoda.A., Taylor.T.D., Hong-Seog.P.,
            Fujiyama.A., Yada.T., Totoki.Y., Watanabe.H. and Sakaki.Y.
            Direct Submission
            Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
            On Jul 13, 2001 this sequence version replaced gi:11995047.
            Location/Qualifiers
            1..176292
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="11"
                /map="11q"
                /clone="RP11-767M8"

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT:
FEATURES
Source
    1..176292
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="11"
        /map="11q"
        /clone="RP11-767M8"

```

```

BASE COUNT      53770 a 32138 c 32968 g 57416 t
ORIGIN
Query Match      1.8%; Score 25; DB 9; Length 176292;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTTCAGAAAAAAGAAAAA 1381
      |||||||
Db 99603 GTTTTTCAGAAAAAAGAAAAA 99579

RESULT 33
AC098950
LOCUS      AC098950      178726 bp      DNA      linear      HTG 20-DEC-2001
DEFINITION Rattus norvegicus clone CH230-175D13, *** SEQUENCING IN PROGRESS
             ***. 64 unordered pieces.
ACCESSION  AC098950
VERSION     AC098950.2  GI:17973734
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
ORGANISM    Rattus norvegicus
            Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
            Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae:
            Rattus.
            1 (bases 1 to 178726)
            Muzny.D.M., Adams.C., Adio-Oduola.B., Ali-osman.F.R., Allen.C.,
            Albrooks.S.L., Amaratunge.H.C., Are.J.R., Banks.T., Barbaria.J.,
            Benton.J., Blimage.K., Blankenburg.K., Bonnin.D., Bouck.J.,
            Bowie.S., Brileva.M., Brown.E., Brown.M., Bryant.N.P., Buhay.C.,
            Burch.P., Burkett.C., Burrell.K.L., Byrd.N.C., Carron.T.F.,
            Carter.M., Cavazos.S.R., Chacko.J., Chavez.D., Chen.G., Chen.R.,
            Chen.Z., Chowdhry.L., Christopoulos.C., Cleveland.C.D., Cox.C.,
            Coyle.M.D., Dathorne.S.R., David.R., Davila.M.L., Davis.C.,
            Davy-Carroll.L., Dederich.D.A., Delaney.K.R., Delgado.O.,
            Denn.A.L., Ding.Y., Dinh.H.H., Douthwaite.K.J., Draper.H.,
            Dugan-Rocha.S., Durbin.K.J., Earnhart.C., Edgar.D., Edwards.C.C.,
            Elhaj.C., Escotto.M., Falls.T., Ferraguto.D., Flagg.N., Ford.J.,
            Foster.P., Frantz.P., Gabisi.A., Gao.J., Garcia.A., Garner.T.,
            Garza.N., Gill.R., Gorrell.J.H., Guevara.W., Gunaratne.P., Hale.S.,
            Hamilton.K., Harris.C., Harris.K., Hart.M., Havlak.P., Hawes.A.,
            Hernandez.J., Hernandez.O., Hodgson.A., Hogues.M., Holloway.C.,
            Hollins.B., Homsli.F., Howard.S., Huber.J., Huiyk.S., Hume.J.,
            Jackson.L.E., Jacobson.B., Jia.Y., Johnson.R., Jolivet.S.,
            Joudah.S., Karlsson.E., Kelly.S., Khan.U., King.L., Korvah.J.,
            Kovar.C., Kratovic.J., Kureshi.A., Landry.N., Leal.B., Lewis.L.C.,
            Lewis.L., Li.J., Li.Z., Lichtarge.O., Lieu.C., Liu.J., Liu.W.,
            Loulseged.H., Lozado.R.J., Lu.X., Lucier.A., Lucier.R., Luna.R.,
            Ma.J., Maheshwari.M., Mapua.P., Martin.R., Martindale.A.,
            Martinez.E., Massey.E., Mawhiney.E., McLeod.M.P., Meador.M.,
            Mei.G., Metzker.M., Miner.G., Miner.Z., Mitchell.T., Mohabbat.K.,
            Morgan.M., Morris.S., Moser.M., Neal.D., Newton.J., Newton.N.,
            Nguyen.A., Nguyen.N., Nguyen.N., Nickerson.E., Nwokenkwo.S.,
            Oguh.M., Okwuonu.G., Oragunye.N., Oviedo.R., Pace.A., Payton.B.,
            Peery.J., Perez.L., Peters.L., Pickens.R., Primus.E., Pu.L.L.,
            Quiles.M., Ren.Y., Rives.M., Rojas.A., Rojibokan.I., Rolfe.M.,
            Ruiz.S., Savary.G., Scherer.S., Scott.G., Shen.H., Shoostari.N.,
            Sison.I., Sodergren.E., Sonaik.T., Sparks.A., Stanley.H.,
            Stone.H., Sutton.A., Svatek.A., Tabor.P., Tamerisa.A., Tamerisa.K.,
            Tang.H., Tansey.J., Taylor.C., Taylor.T., Telford.B., Thomas.N.,
            Thomas.S., Usmani.K., Vasquez.L., Vera.V., Villalon.D., Vinson.R.,
            Wall.R., Wang.S., Ward-Moore.S., Warren.R., Washington.C.,
            Watlington.S., Williams.G., Williamson.A., Wleczyk.R., Wooden.S.,
            Worley.K., Wu.C., Wu.Y., Wu.Y.F., Zhou.J., Zorrilla.S., Nelson.D.,
            Weinstock.G. and Gibbs.R.
            Direct Submission
            Unpublished
            2 (bases 1 to 178726)
            Worley.K.C.
            Direct Submission
            Submitted (07-NOV-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA

```

COMMENT

On Dec 20, 2001 this sequence version replaced gi:16756144.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GJF
 Center clone name: CH230-175D13
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329First call to
 findPhrapList
 Consensus quality: 142246 bases at least Q40
 Consensus quality: 149877 bases at least Q30
 Consensus quality: 156651 bases at least Q20
 Estimated insert size: 147873; sum-of-contigs estimation
 Quality coverage: 0x in Q20 bases; agarose-fp estimation
 Quality coverage: 2.2x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 64 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence,
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 7357: contig of 7357 bp in length
 * 7358 7457: gap of unknown length
 * 7458 15306: contig of 7849 bp in length
 * 15307 15406: gap of unknown length
 * 15407 24299: contig of 8893 bp in length
 * 24300 24399: gap of unknown length
 * 24400 30138: contig of 5739 bp in length
 * 30139 30238: gap of unknown length
 * 30239 36252: contig of 6014 bp in length
 * 36253 36352: gap of unknown length
 * 36353 42882: contig of 5930 bp in length
 * 42883 42882: gap of unknown length
 * 42883 47458: contig of 5076 bp in length
 * 47459 47558: gap of unknown length
 * 47559 52616: contig of 5058 bp in length
 * 52617 52716: gap of unknown length
 * 52717 57624: contig of 4908 bp in length
 * 57625 57724: gap of unknown length
 * 57725 62717: contig of 4993 bp in length
 * 62718 62817: gap of unknown length
 * 62818 65192: contig of 2375 bp in length
 * 65193 65292: gap of unknown length
 * 65293 68632: contig of 3340 bp in length
 * 68633 68732: gap of unknown length
 * 68733 72375: contig of 3643 bp in length
 * 72376 72476: gap of unknown length
 * 72476 75816: contig of 3341 bp in length
 * 75817 75916: gap of unknown length
 * 75917 79213: contig of 3297 bp in length
 * 79214 79313: gap of unknown length
 * 79314 82693: contig of 3380 bp in length
 * 82694 82793: gap of unknown length
 * 82794 85333: contig of 2540 bp in length
 * 85334 85433: gap of unknown length
 * 85434 89418: contig of 3985 bp in length
 * 89419 89518: gap of unknown length
 * 89519 93053: contig of 3535 bp in length
 * 93054 93153: gap of unknown length
 * 93154 95888: contig of 2735 bp in length
 * 95889 95988: gap of unknown length
 * 95989 99179: contig of 3191 bp in length
 * 99180 99279: gap of unknown length
 * 99280 101034: contig of 1755 bp in length
 * 101035 101134: gap of unknown length

101135 103242: contig of 2108 bp in length
 * 103243 103342: gap of unknown length
 * 103343 105210: contig of 1868 bp in length
 * 105211 105310: gap of unknown length
 * 105311 107490: contig of 2180 bp in length
 * 107491 107590: gap of unknown length
 * 107591 110892: contig of 3302 bp in length
 * 110893 110992: gap of unknown length
 * 110993 113287: contig of 2295 bp in length
 * 113288 113387: gap of unknown length
 * 113388 116067: contig of 2680 bp in length
 * 116068 116167: gap of unknown length
 * 116168 118724: contig of 2557 bp in length
 * 118725 118824: gap of unknown length
 * 118825 120414: contig of 1590 bp in length
 * 120415 120514: gap of unknown length
 * 120515 121596: contig of 1082 bp in length
 * 121597 121696: gap of unknown length
 * 121697 124109: contig of 2413 bp in length
 * 124110 124209: gap of unknown length
 * 124210 126884: contig of 2655 bp in length
 * 126885 126984: gap of unknown length
 * 126985 129136: contig of 2172 bp in length
 * 129137 129236: gap of unknown length
 * 129237 131988: contig of 2752 bp in length
 * 131989 132088: gap of unknown length
 * 132089 134242: contig of 2154 bp in length
 * 134243 136424: gap of unknown length
 * 136425 138196: contig of 1673 bp in length
 * 138197 138296: gap of unknown length
 * 138297 139992: contig of 1696 bp in length
 * 139993 140092: gap of unknown length
 * 140093 141978: contig of 1886 bp in length
 * 141979 142078: gap of unknown length
 * 142079 143585: contig of 1487 bp in length
 * 143586 143685: gap of unknown length
 * 143686 145077: contig of 1412 bp in length
 * 145078 145177: gap of unknown length
 * 145178 147331: contig of 2154 bp in length
 * 147332 147431: gap of unknown length
 * 147432 148677: contig of 1246 bp in length
 * 148678 148777: gap of unknown length
 * 148778 150439: contig of 1662 bp in length
 * 150440 150539: gap of unknown length
 * 150540 151753: contig of 1214 bp in length
 * 151754 151853: gap of unknown length
 * 151854 154142: contig of 2289 bp in length
 * 154143 154242: gap of unknown length
 * 154243 155932: contig of 1690 bp in length
 * 155933 156032: gap of unknown length
 * 156033 157178: contig of 1146 bp in length
 * 157179 157278: gap of unknown length
 * 157279 158371: contig of 1093 bp in length
 * 158372 158471: gap of unknown length
 * 158472 159477: contig of 1006 bp in length
 * 159478 159577: gap of unknown length
 * 159578 161073: contig of 1496 bp in length
 * 161074 161173: gap of unknown length
 * 161174 162805: contig of 1632 bp in length
 * 162806 162905: gap of unknown length
 * 162906 164379: contig of 1474 bp in length
 * 164380 164479: gap of unknown length
 * 164480 166196: contig of 1717 bp in length
 * 166197 166296: gap of unknown length
 * 166297 167647: contig of 1351 bp in length
 * 167648 167747: gap of unknown length

Query Match 1.8%; Score 25; DB 2; Length 178726;
 Best Local Similarity 100.0%; Pred. No. 0.021;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

QY 1357 GTTTTGAAGAAAAAAGAAAAA 1381
      |||||||
Db 162064 GTTTTGAAGAAAAAAGAAAAA 162088
      ***, 66 unordered pieces.

AC105354      182365 bp      DNA      linear      HTG 02-JAN-2002
Rattus norvegicus clone CH230-97J10, *** SEQUENCING IN PROGRESS

AC105354
AC105354.1 GI:18034693
HTG: HTGS_PHASE1.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 182365)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Aisbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
Coyle,M.D., Dathorne,S.R., David,R., Davila,M.I., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Duhan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louisleg,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogulu,M., Okwundu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N.,
Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H.,
Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
Worley,K., Wu,C., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 182365)
Worley,K.C.
Direct Submission
Submitted (02-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
----- Center project name: G1XR

```

```

Center clone name: CH230-97J10
----- Summary Statistics
Assembly program: Phrap; version 0.990329First call to
findphraplist
Consensus quality: 145810 bases at least Q40
Consensus quality: 152522 bases at least Q30
Consensus quality: 158737 bases at least Q20
Estimated insert size: 158711; sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 2.5x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 66 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 6187: contig of 6187 bp in length
* 6287: gap of unknown length
* 6288
* 11317: contig of 5030 bp in length
* 11318
* 11417: gap of unknown length
* 11418
* 13221: contig of 7804 bp in length
* 19222
* 19321: gap of unknown length
* 19322
* 27390: contig of 8069 bp in length
* 27391
* 27490: gap of unknown length
* 27491
* 3712: contig of 6222 bp in length
* 3713
* 3812: gap of unknown length
* 3813
* 38766: contig of 4954 bp in length
* 38767
* 38866: gap of unknown length
* 38867
* 4444: contig of 5578 bp in length
* 4445
* 4454: gap of unknown length
* 4455
* 49828: contig of 5284 bp in length
* 49829
* 49928: gap of unknown length
* 49929
* 54540: contig of 4612 bp in length
* 54541
* 54640: gap of unknown length
* 54641
* 58440: contig of 3800 bp in length
* 58441
* 58540: gap of unknown length
* 58541
* 62837: contig of 4297 bp in length
* 62838
* 62937: gap of unknown length
* 62938
* 66668: contig of 3731 bp in length
* 66669
* 66768: gap of unknown length
* 66769
* 70974: contig of 4206 bp in length
* 70975
* 71074: gap of unknown length
* 71075
* 74368: contig of 3294 bp in length
* 74369
* 74668: gap of unknown length
* 74669
* 77982: contig of 3514 bp in length
* 77983
* 78082: gap of unknown length
* 78083
* 81004: contig of 2922 bp in length
* 81005
* 81104: gap of unknown length
* 81105
* 84555: contig of 3351 bp in length
* 84556
* 84555: gap of unknown length
* 84556
* 87164: contig of 2609 bp in length
* 87165
* 87264: gap of unknown length
* 87265
* 91544: contig of 4280 bp in length
* 91545
* 91644: gap of unknown length
* 91645
* 94067: contig of 2423 bp in length
* 94068
* 94167: gap of unknown length
* 94168
* 97458: contig of 3291 bp in length
* 97459
* 97558: gap of unknown length
* 97559
* 100659: contig of 3101 bp in length
* 100660
* 100759: gap of unknown length
* 100760
* 103998: contig of 3239 bp in length
* 103999
* 104098: gap of unknown length
* 104099
* 106554: contig of 2456 bp in length
* 106555
* 106554: gap of unknown length
* 106555
* 108753: contig of 2099 bp in length
* 108754
* 108853: gap of unknown length
* 112113: contig of 3260 bp in length
* 112114
* 112213: gap of unknown length

```

AC019108 185721 bp DNA linear HTG 07-JUL-2000
Homo sapiens chromosome 11 clone RP11-485A5, WORKING DRAFT
SEQUENCE, 32 unordered pieces.
AC019108
AC019108.5 GI:8570235
HTG; HTGS_PHASE1; HTGS_DRAFT.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston, R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 185721)
Waterston, R.H.
Direct Submission
Submitted (30-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jun 17, 2000 this sequence version replaced gi:7232162.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0485A05
----- Summary Statistics -----
Sequencing vector: M13; 70%
Sequencing vector: plasmid; 30%
Chemistry: Dye-primer ET; 70% of reads
Chemistry: Dye-terminator Big Dye; 30% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 168100 bases at least Q40
Consensus quality: 173526 bases at least Q30
Consensus quality: 176219 bases at least Q20
Insert size: 181000; agarose-fp
Insert size: 182621; sum-of-contigs
Quality coverage: 3.59 in Q20 bases; agarose-fp
Quality coverage: 3.52 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1227: contig of 1227 bp in length
* 1228: gap of unknown length
* 1328: contig of 1419 bp in length
* 2747: gap of unknown length
* 2846: contig of 1639 bp in length
* 4485: gap of unknown length
* 4585: gap of unknown length
* 6449: contig of 1864 bp in length
* 6549: gap of unknown length
* 7840: contig of 1291 bp in length
* 7841: gap of unknown length
* 9479: contig of 1539 bp in length
* 9579: gap of unknown length
* 9580: contig of 1404 bp in length
* 10983: gap of unknown length
* 11084: contig of 1314 bp in length
* 12398: gap of unknown length
* 12498: contig of 2303 bp in length
* 14801: gap of unknown length
* 14901: contig of 1958 bp in length
* 16859: gap of unknown length
* 16959: contig of 1895 bp in length
* 18854: gap of unknown length
* 18954: contig of 1622 bp in length

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
COMMENT

114311: contig of 2098 bp in length
114312: gap of unknown length
116394: contig of 1983 bp in length
116494: gap of unknown length
119820: contig of 3326 bp in length
119920: gap of unknown length
123268: contig of 3348 bp in length
123668: gap of unknown length
124983: contig of 1615 bp in length
125083: gap of unknown length
127840: contig of 2757 bp in length
127940: gap of unknown length
129294: contig of 1354 bp in length
129394: gap of unknown length
131286: contig of 1892 bp in length
131386: gap of unknown length
131387: contig of 1341 bp in length
132827: gap of unknown length
134355: contig of 1528 bp in length
134455: gap of unknown length
136447: contig of 1992 bp in length
136547: gap of unknown length
137913: contig of 1366 bp in length
138013: gap of unknown length
139576: contig of 1563 bp in length
139577: gap of unknown length
141967: contig of 2291 bp in length
142067: gap of unknown length
143845: contig of 1778 bp in length
143846: gap of unknown length
143945: gap of unknown length
145316: contig of 1371 bp in length
145416: gap of unknown length
147462: contig of 2046 bp in length
147562: gap of unknown length
149333: contig of 1771 bp in length
149433: gap of unknown length
150802: contig of 1369 bp in length
150902: gap of unknown length
152836: contig of 1934 bp in length
152936: gap of unknown length
154375: contig of 1239 bp in length
154775: gap of unknown length
155922: contig of 1647 bp in length
156022: gap of unknown length
157163: contig of 1141 bp in length
157263: gap of unknown length
158812: contig of 1549 bp in length
158912: gap of unknown length
160656: contig of 1744 bp in length
160756: gap of unknown length
162544: contig of 1788 bp in length
162644: gap of unknown length
163846: contig of 1202 bp in length
163946: gap of unknown length
165035: contig of 1089 bp in length
165135: gap of unknown length
165136: contig of 1453 bp in length
166588: gap of unknown length
166889: contig of 1573 bp in length
168361: gap of unknown length
169900: contig of 1539 bp in length

Query Match 1.8%; Score 25; DB 2; Length 182365;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1355 CTGTTTTTGAAGAAAAA 1379
|||||TTTTTTTTTTTTTTTT
Db 154840 CTGTTTTTGAAGAAAAA 154816

RESULT 35
AC019108

```
* 20576 20675: gap of unknown length
* 20676 22301: contig of 1626 bp in length
* 22302 22401: gap of unknown length
* 22402 24505: contig of 2104 bp in length
* 24506 24605: gap of unknown length
* 24606 26368: contig of 1763 bp in length
* 26369 26468: gap of unknown length
* 26469 29075: contig of 2607 bp in length
* 29076 32175: gap of unknown length
* 32176 32702: contig of 3527 bp in length
* 32703 32802: gap of unknown length
* 32803 36360: contig of 3558 bp in length
* 36361 36460: gap of unknown length
* 36461 39607: contig of 3146 bp in length
* 39608 43982: contig of 4276 bp in length
* 43983 44082: gap of unknown length
* 44083 49628: contig of 5546 bp in length
* 49629 49728: gap of unknown length
* 49729 55205: contig of 5477 bp in length
* 55206 55305: gap of unknown length
* 55306 63685: contig of 8380 bp in length
* 63686 63785: gap of unknown length
* 63786 70762: contig of 6977 bp in length
* 70763 70862: gap of unknown length
* 70863 78073: contig of 7217 bp in length
* 78074 78173: gap of unknown length
* 78174 86155: contig of 7976 bp in length
* 86156 86255: gap of unknown length
* 86256 95500: contig of 9245 bp in length
* 95501 95600: gap of unknown length
* 95601 103045: contig of 7446 bp in length
* 103046 103145: gap of unknown length
* 103146 111155: contig of 8009 bp in length
* 111156 111255: gap of unknown length
* 111256 133307: contig of 22052 bp in length
* 133308 133407: gap of unknown length
* 133408 159247: contig of 25840 bp in length
* 159248 159347: gap of unknown length
* 159348 185721: contig of 26374 bp in length.

FEATURES             Location/Qualifiers
     source
     ..185721
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /chromosome="11"
     /clone="RP11-485A5"

BASE COUNT  58471 a 33336 c 34184 g 56600 t 3130 others
ORIGIN

Query Match          1.8%; Score 25; DB 2; Length 185721;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTGAAGAAAAAAGAAAAA 1381
      |||||
Db 17365 GTTTTGAAGAAAAAAGAAAAA 17389

RESULT 36
AX336388          267156 bp   DNA   linear   PAT 09--JAN-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horriqan, S., Soppet, D.R. and Weaver, Z.
Cancer gene determination and therapeutic screening using signature

gene sets
Patent: WO 0194629-A 6897 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
     Location/Qualifiers
     ..267156
     /organism="Homo sapiens"
     /db_xref="taxon:9606"

BASE COUNT  76527 a 56343 c 55787 g 78499 t
ORIGIN

Query Match          1.8%; Score 25; DB 6; Length 267156;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTGAAGAAAAAAGAAAAA 1381
      |||||
Db 68711 GTTTTGAAGAAAAAAGAAAAA 68735

RESULT 37
U66059          267156 bp   DNA   linear   PRI 30-NOV-1998
LOCUS
DEFINITION
Humau germline T-cell receptor beta chain
Dopamine-beta-hydroxylase-like, TRV1, TRY2, TRY3, TCRBV2S1P,
TCRBV22S1A2N1T, TCRBV9S1A1T, TCRBV7S1A1N2T, TCRBV5S1A1T, TCRBV13S3,
TCRBV6S7P, TCRBV7S3A2T, TCRBV13S2A1T, TCRBV9S2A2PT, TCRBV7S2A1N4T,
TCRBV13S9/13S2A1T, TCRBV6S5A1N1, TCRBV30S1P, TCRBV31S1, TCRBV13S5,
TCRBV6S1A1N1, TCRBV32S1P, TCRBV5S5P, TCRBV1S1A1N1, TCRBV12S2A1T,
TCRBV21S1, TCRBV8S4P, TCRBV12S3, TCRBV21S3A2N2T, TCRBV8S5P,
TCRBV13S1 genes from bases 1 to 267156 (section 1 of 3).
U66059 L36092
U66059.1 GI:1552494
C region: C-beta gene segment; D region: J-beta gene segment;
J-segment: T-cell receptor beta-chain; TCR-beta gene; V-beta gene
segment; V-segment: cell membrane protein; constant region;
diversity region; germline: joining segment; trypsin; trypsinogen;
variable segment.
Homo sapiens (clone: K41A) (clone_lib: CGM1: YAC D49H4) DNA; Homo
sapiens (clone: K35) (clone_lib: CGM1: YAC D49H4) DNA; Homo sapiens
(clone: K26) (clone_lib: CGM1: YAC D49H4) DNA; Homo sapiens (clone:
K56) (clone_lib: CGM1: YAC D49H4) DNA; Homo sapiens (library:
ATCC1475) (clone: X21B) (clone_lib: Kai Wang's) DNA; Homo sapiens
(library: ATCC1475) (clone: G54) (clone_lib: Kai Wang's) DNA; Homo
sapiens (library: HeLa) (clone: H137) (clone_lib: Eric Lai's) DNA;
Homo sapiens (library: sperm) (clone: H18) (clone_lib: Eric Lai's)
DNA; Homo sapiens (clone: H18/G15gap) DNA; Homo sapiens (library:
ATCC1475) (clone: G15) (clone_lib: Kai Wang's) DNA; Homo sapiens
(library: ATCC1475) (clone: X1A) (clone_lib: Kai Wang's) DNA; Homo
sapiens (clone: A27) (clone_lib: CGM1: YAC 234 A6F6) DNA; Homo
sapiens (clone: A212partial) (clone_lib: CGM1: YAC 234 A6F6) DNA;
Homo sapiens (clone: A14) (clone_lib: CGM1: YAC 234 A6F6) DNA; Homo
sapiens (library: sperm) (clone: H7.1) (clone_lib: Eric Lai's) DNA;
Homo sapiens (clone: H12.18) (clone_lib: Eric Lai's) DNA; Homo
sapiens (library: HeLa) (clone: H130) (clone_lib: Eric Lai's) DNA;
Homo sapiens (clone: A16) (clone_lib: CGM1: YAC 234 A6F6) DNA; Homo
sapiens (clone: C215) (clone_lib: CGM1: YAC 234 A7B3) DNA; Homo
sapiens (clone: G1) (library: ATCC1475) (clone_lib: Kai Wang's)
DNA; Homo sapiens (clone: C68) (clone_lib: CGM1: YAC 210 A38G1)
DNA; Homo sapiens (clone: C21) (clone_lib: CGM1: YAC 234 A72B3)
DNA; Homo sapiens (library: ATCC1475) (clone: X11) (clone_lib: Kai
Wang's) DNA; Homo sapiens (library: ATCC1475) (clone: X6A)
(clone_lib: Kai Wang's) DNA; Homo sapiens (library: ATCC1475)
(clone: CBG1) (clone_lib: Kai Wang's) DNA; Homo sapiens (clone:
CBG1/C29gap) DNA; and Homo sapiens (library: ATCC1475) (clone: C29)
(clone_lib: Kai Wang's) DNA.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 267156)
Rowen, L., Koop, B.F. and Hood, L.
The complete 685-kilobase DNA sequence of the human beta T cell
receptor locus
Science 272 (5269), 1755-1762 (1996)
JOURNAL
```



```

exon      /rpt_family="MER42C"
complement(33276..33406)

Query Match
Best Local Similarity 1.8%; Score 25; DB 9; Length 267156;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTGTGAAAAAAGAAAAA 1381
|||||
Db 68711 GTTTTGTGAAAAAAGAAAAA 68735

RESULT 38
AF124738
LOCUS      Zea mays clone CODAM33 unknown mRNA.          PLN 29-JUN-1999
DEFINITION
ACCESSION  AF124738
VERSION    AF124738.1 GI:5257287
KEYWORDS   Zea mays.
SOURCE     Zea mays.
ORGANISM   Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 230)
Tikhonov,A.P., SanMiguel,P.J., Nakajima,Y., Gorenstein,N.M.,
Bennetzen,J.L. and Avramova,Z.
Colinearity and its exceptions in orthologous adh regions of maize
and sorghum
Proc. Natl. Acad. Sci. U.S.A. 96 (13), 7409-7414 (1999)
2 (bases 1 to 230)
Tikhonov,A.P., SanMiguel,P.J., Nakajima,Y., Gorenstein,N.M.,
Bennetzen,J.L. and Avramova,Z.
Direct Submission
Submitted (01-FEB-1999) of Biological Sciences, Purdue University,
1392 Lilly Hall, West Lafayette, IN 47907, USA
FEATURES
Location/Qualifiers
1..230
/organism="Zea mays"
/variety="B73"
/db_xref="taxon:4577"
/clone="CODAM33"
/tissue_type="whole seedling"
/clone_lib="Clontech #FL1032n"
/dev_stage="2 leaf stage"
/note="extrachromosomal"
<1..147
/note="similar to maize 334B7.1 gene and sorghum 110K5.1
gene"
/codon_start=1
/evidence-not_experimental
/product="unknown"
/protein_id="AADA1268.1"
/db_xref="GI:5257288"
/translation="TSLKHLVKILDQAPAKRSICGTPKSIIVTLFMVCRHRTGIDHFL
LCWL"
BASE COUNT      80 a 33 c 48 g 69 t
ORIGIN

Query Match
Best Local Similarity 1.7%; Score 24; DB 8; Length 230;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAAAGAAAAA 1381
|||||
Db 204 TTTTGTGAAAAAAGAAAAA 227

RESULT 39
AX314458/c
LOCUS      AX314458
DEFINITION  Sequence 7443 from Patent WO0190366.
ACCESSION  AX314458

Query Match
Best Local Similarity 1.7%; Score 24; DB 6; Length 387;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAAAGAAAAA 1381
|||||
Db 69 TTTTGTGAAAAAAGAAAAA 46

RESULT 41
AX186327
LOCUS      AX186327
DEFINITION  Sequence 2022 from Patent WO0142467.
ACCESSION  AX186327
VERSION    AX186327.1 GI:15137759
KEYWORDS   human.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

VERSION     AX314458.1 GI:17898464
KEYWORDS    human.
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
AUTHORS     Leach,M.D. and Shinkets,R.A.
TITLE       Human polynucleotides and polypeptides encoded thereby
JOURNAL     Patent: WO 0190366-A 7443 29-NOV-2001;
            Curagen Corporation (US)
FEATURES
Source      Location/Qualifiers
1..292
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT  63 a 71 c 46 g 112 t
ORIGIN

Query Match
Best Local Similarity 1.7%; Score 24; DB 6; Length 292;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAAAGAAAAA 1381
|||||
Db 148 TTTTGTGAAAAAAGAAAAA 125

RESULT 40
AX069660/c
LOCUS      AX069660
DEFINITION  Sequence 132 from Patent WO0102568.
ACCESSION  AX069660
VERSION    AX069660.1 GI:12579445
KEYWORDS   human.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 387)
AUTHORS     Williams,L.T., Escobedo,J., Innis,M.A., Garcia,P.D., Klinger,J.,
            Kassam,A., Relhan,C., Randazzo,F., Kennedy,G.C., Pot,D.,
            Lamson,G., Drmanac,R., Crkenjakov,R., Drmanac,S., Dickson,M.,
            Labat,I., Leshkowitz,D., Kita,D., Garcia,V. and Strache-Crain,B.
TITLE       Human genes and gene expression products
JOURNAL     Patent: WO 0102568-A 132 11-JAN-2001;
            CHIRON CORPORATION (US); HYSEQ, INC. (US)
FEATURES
Source      Location/Qualifiers
1..387
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT  127 a 63 c 97 g 100 t
ORIGIN

Query Match
Best Local Similarity 1.7%; Score 24; DB 6; Length 387;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAAAGAAAAA 1381
|||||
Db 69 TTTTGTGAAAAAAGAAAAA 46

RESULT 41
AX186327
LOCUS      AX186327
DEFINITION  Sequence 2022 from Patent WO0142467.
ACCESSION  AX186327
VERSION    AX186327.1 GI:15137759
KEYWORDS   human.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 573)
AUTHORS Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
TITLE Genes, compositions, kits, and methods for identification,
JOURNAL assessment, prevention, and therapy of cervical cancer
Patent: WO 0142467-A 2022 14-JUN-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source Location/Qualifiers
1..573
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 161 a 121 c 133 g 158 t
ORIGIN
Query Match 1.7%; Score 24; DB 6; Length 573;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1358 TTTTGTGAAAAA...AAAAAAAAA 1381
|||||
Db 517 TTTTGTGAAAAA...AAAAAAAAA 540

RESULT 42
AX187052 597 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 2747 from Patent WO0142467.
ACCESSION AX187052
VERSION AX187052.1 GI:15138496
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 597)
AUTHORS Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
TITLE Genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of cervical cancer
JOURNAL Patent: WO 0142467-A 2747 14-JUN-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source Location/Qualifiers
1..597
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 259 a 54 c 39 g 244 t
ORIGIN
Query Match 1.7%; Score 24; DB 6; Length 597;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1358 TTTTGTGAAAAA...AAAAAAAAA 1381
|||||
Db 310 TTTTGTGAAAAA...AAAAAAAAA 333

RESULT 43
S82852 648 bp mRNA linear ROD 11-FEB-1997
LOCUS tmk-thymidylate kinase {alternatively spliced} [mice, BALB/c, 3T3]
DEFINITION cells mRNA, 648 nt.
ACCESSION S82852
VERSION S82852.1 GI:1836039
KEYWORDS
SOURCE
MUS sp. BALB/c 3T3 cells.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 648)
AUTHORS Liang,P., Averbouk,L., Zhu,W., Haley,T. and Pardee,A.B.
TITLE Molecular characterization of the murine thymidylate kinase gene
JOURNAL Cell Growth Differ. 6 (10), 1333-1338 (1995)
MEDLINE 96114098
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gbbseq 179234] from the original journal article.
This sequence comes from Fig. 2A.
FEATURES
source Location/Qualifiers
1..648
/organism="Mus sp."
/db_xref="taxon:10095"
1..648
/feature="tmk"
/feature="thymidylate kinase, TMK"
49..291
/feature="tmk"
/feature="catalyzes conversion of dTMP to dTDP; This sequence
comes from Fig. 2A; TMK"
/codon_start=1
/product="thymidylate kinase"
/protein_id="AA046837.1"
/db_xref="GI:1836040"
/translation="MASRRGALIVLEGVDRAKTTQGLKLVLTALCASGHRAELLRFPE
RSTEIGKLLNSYLEKTELEDSHVLFSANWEQV"
BASE COUNT 197 a 139 c 163 g 149 t
ORIGIN
Query Match 1.7%; Score 24; DB 10; Length 648;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1358 TTTTGTGAAAAA...AAAAAAAAA 1381
|||||
Db 586 TTTTGTGAAAAA...AAAAAAAAA 609

RESULT 44
DDIUPDPGP/c 700 bp DNA linear INV 29-DEC-1995
LOCUS Dictyostelium discoideum UDP-glucose pyrophosphorylase (UDPGP1)
DEFINITION gene, 5' flank.
ACCESSION M27639.1 GI:341741
VERSION M27639.1
KEYWORDS UDP-glucose pyrophosphorylase.
SOURCE Dictyostelium discoideum DNA.
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE 1 (bases 1 to 700)
AUTHORS Pavlovic,J., Haribabu,B. and Dottin,R.P.
TITLE Transmembrane signal transduction regulates gene expression in
JOURNAL Dictyostelium discoideum
MEDLINE Dev. Genet. 9 (4-5), 371-382 (1988)
FEATURES
source Location/Qualifiers
1..700
/organism="Dictyostelium discoideum"
/db_xref="taxon:44689"
610..700
/feature="UDPGP1"
610..700
/feature="UDPGP1"
/codon_start=1
/product="UDP-glucose pyrophosphorylase"
/protein_id="AA084450.1"
/db_xref="GI:1141709"
/translation="MTDTATSKATVERPKLOSTGSLHSLFKDND"
BASE COUNT 288 a 57 c 36 g 319 t
ORIGIN
Query Match 1.7%; Score 24; DB 3; Length 700;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1358 TTTTGTGAAAAA...AAAAAAAAA 1381
|||||
Db 267 TTTTGTGAAAAA...AAAAAAAAA 244

```

```

RESULT 45
AF339830
LOCUS Homo sapiens clone IMAGE:609899, mRNA sequence.
DEFINITION AF339830
ACCESSION AF339830
VERSION AF339830.1 GI:13507368
KEYWORDS FLI_CDNA.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 766)
Christian,S.L., Vlamakis,V., Liu,C.-Y., Badner,J.A. and
Gershon,E.S.
Integrated transcript map including 65 sequenced cDNA clones for
the 13q32 region linked to both bipolar disorder and schizophrenia
2 (bases 1 to 766)
Unpublished
Christian,S.L., Vlamakis,V., Liu,C.-Y., Badner,J.A. and
Gershon,E.S.
Direct Submission
TITLE Submitted (22-JAN-2001) Psychiatry, University of Chicago, 924 East
JOURNAL 57th Street, Room R014, Chicago, IL 60637-1455, USA
FEATURES
Source
1..766
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="13"
/map="13q32"
/clone="IMAGE:609899"
/clone_lib="Stratagene fetal retina #937202"
BASE COUNT 202 a 147 c 171 g 246 t
ORIGIN

Query Match 1.7%; Score 24; DB 9; Length 766;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAAAAAAAAAAAAAA 1381
|||||
Db 742 TTTTGTGAAAAAAAAAAAAAAAAA 765
|||||

```

Search completed: November 5, 2002, 17:35:27
 Job time : 7490.33 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 10:52:52 ; Search time 164.876 Seconds
(without alignments)
14380.822 Million cell updates/sec

Title: US-09-805-311-5
Perfect score: 1381
Sequence: 1 cgaccacacgctcgccac.....ttgaaaaaaaaaaaaaaaa 1381

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_032802.*

```

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1381	100.0	1381	21 AAA27925	Maize Rad2/FEN-1 c
2	1267	91.7	1463	21 AAA27923	Maize Rad2/FEN-1 c
3	1216	88.1	1541	21 AAA27924	Maize Rad2/FEN-1 c
4	850	61.5	1478	21 AAA27926	Maize Rad2/FEN-1 c
5	27	2.0	1096	20 AAX23270	A. niger 2,3-dihyd
6	25	1.8	392	22 AA185371	Human polynucleoti
7	25	1.8	1255	21 AA185371	Human secreted pro
8	25	1.8	1570	21 AAC76262	Human ORFX ORF1817
9	25	1.8	1648	21 AAC59600	Human secreted pro

c 11	24	1.7	343	22 AAD05605	Human secreted pro
c 12	24	1.7	387	22 AAF64376	Novel human polynu
c 13	24	1.7	403	22 AAI82336	Human polynucleoti
14	24	1.7	406	22 AAI89042	Human polynucleoti
15	24	1.7	425	22 AAK58115	Human immune/haema
16	24	1.7	425	22 AAI92294	Human polynucleoti
17	24	1.7	469	22 AAK66234	Human immune/haema
18	24	1.7	500	21 AAC78423	Human cancer assoc
19	24	1.7	573	22 AAH70748	Human cervical can
20	24	1.7	579	22 AAD05657	Human secreted pro
21	24	1.7	597	22 AAH71472	Human secreted pro
22	24	1.7	860	22 AAK71980	Human immune/haema
c 23	24	1.7	860	22 AAK71981	Human immune/haema
24	24	1.7	1061	21 AAF09266	Fusarium venenatum
25	24	1.7	1304	19 AAV30555	Strawberry chalcon
26	24	1.7	1311	22 AAF84329	Barley iron absorp
27	24	1.7	1312	22 AAF84327	Barley iron absorp
28	24	1.7	1560	21 AAK79701	Human secreted pro
29	24	1.7	1648	21 AAC64785	Lemon acyl transfe
30	24	1.7	1727	21 AAC81068	Human secreted pro
31	24	1.7	1807	22 AAD15431	Wheat protoporphy
32	24	1.7	1811	18 AAV04308	Wheat protox-1 cdn
33	24	1.7	1811	18 AAT86121	Wheat protoporphy
34	24	1.7	1811	22 AAF76575	Wheat protoporphy
35	24	1.7	1819	21 AAC59348	Human secreted pro
36	24	1.7	2010	21 AAC59604	Human secreted pro
37	24	1.7	2080	21 AAC25264	Human secreted pro
38	24	1.7	2082	19 AAV59751	Human secreted pro
39	24	1.7	3350	22 AAD08345	Human secreted pro
c 40	24	1.7	5828	22 ABA19864	Human nervous syst
41	24	1.7	5843	22 ABA19329	Human nervous syst
42	24	1.7	9928	22 ABA19865	Human nervous syst
43	24	1.7	11729	22 ABL32894	Human immune syst
44	24	1.7	3223	22 ABA17368	Human nervous syst
45	24	1.7	3223	22 ABA19863	Human nervous syst

ALIGNMENTS

```

RESULT 1
AAA27925
ID AAA27925 standard; cDNA; 1381 BP.
XX
AC AAA27925;
XX
DT 12-SEP-2000 (first entry)
XX
DE Maize Rad2/FEN-1 cDNA.
XX
KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
KW endonuclease; exonuclease; DNA repair; gene targeting; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 37..1176
FT /*tag= a
XX
PN WO200036109-A1.
XX
PD 22-JUN-2000.
XX
PF 16-NOV-1999; 99WO-US27147.
XX
PR 15-DEC-1998; 98US-0112332.
XX
PA (PION-) PIONEER HI-BRED INT INC.
PI Mahajan PB;
XX
WPI 2000-452026/39.
DR

```

DR P-PSDB; AAY95309.
XX Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT recombination and repair in transgenic plants, e.g. for gene targeting
PT and the production of male sterile plants -
XX Example 1; Page 75-76; 85pp; English.
XX
CC The present sequence is that of maize cDNA coding for RAD2/FEN-1
CC (see AAY95309). The cDNA was isolated from a cDNA library prepared
CC from W23 line tassel (16-18 cm long) polyA RNA. Rad2/FEN-1 is a
CC structure specific endonuclease which under certain conditions also
CC acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to
CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
CC The protein is involved in the regulation of DNA repair and
CC recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.
XX
SQ Sequence 1381 BP; 441 A; 269 C; 346 G; 325 T; 0 other;

Query Match 100.0%; Score 1381; DB 21; Length 1381;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1381; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACCACGCGTCGCGGCACAGCGCGCAGACGAGATGGCATCAAGGTTTGACGAAA 60
DB 1 CGACCACGCGTCGCGGCACAGCGCGCAGACGAGATGGCATCAAGGTTTGACGAAA 60
QY 61 CTGCTGGCGGACATGCGCCCAAGCGCATGAAGAGCAGAAAGTTTCAGAGCTACTTCGCG 120
DB 61 CTGCTGGCGGACATGCGCCCAAGCGCATGAAGAGCAGAAAGTTTCAGAGCTACTTCGCG 120
QY 121 CGCAAAATCGCGCTCGACGCGCAGCATGAGCATATACCACTTCCTGATTTAGTTGAAGG 180
DB 121 CGCAAAATCGCGCTCGACGCGCAGCATGAGCATATACCACTTCCTGATTTAGTTGAAGG 180
QY 181 ACAGCATGGAACCTCTCACAAATGAAGCTGGTGAAGTCACTAGTCATTTCGATGAGG 240
DB 181 ACAGCATGGAACCTCTCACAAATGAAGCTGGTGAAGTCACTAGTCATTTCGATGAGG 240
QY 241 TTCACCGGACAAATGAAGTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGTATGCG 300
DB 241 TTCACCGGACAAATGAAGTACTGGAAGCGGGAATCAAGCCAGTTTATGTTTGTATGCG 300
QY 301 AAGCCTCTGATATGAGAAACAAAGAGCTTGCTAAAAGATACATAAAAGAGATGATGCA 360
DB 301 AAGCCTCTGATATGAGAAACAAAGAGCTTGCTAAAAGATACATAAAAGAGATGATGCA 360
QY 361 ACCAAGATCTGACTAGGCGAGTAGAGTAGAGATAAAGATCGATTTGAAAATTTAGGC 420
DB 361 ACCAAGATCTGACTAGGCGAGTAGAGTAGAGATAAAGATCGATTTGAAAATTTAGGC 420
QY 421 AAGAGGAGCTGTAAGGTCACAAAGCAACACACAGAGATTTGTAACGGCTATTAGACTT 480
DB 421 AAGAGGAGCTGTAAGGTCACAAAGCAACACACAGAGATTTGTAACGGCTATTAGACTT 480
QY 481 ATGGGGTTCCTGTTGAGAGGACCTTCTGAAGCAGACGAATGTCAGCCCTTTCG 540
DB 481 ATGGGGTTCCTGTTGAGAGGACCTTCTGAAGCAGACGAATGTCAGCCCTTTCG 540
QY 541 ATAAACGATAAGGTTTTCGCTGTTGCTTCAGAGATATGGACTCCCTTACTTTTGGGGCT 600
DB 541 ATAAACGATAAGGTTTTCGCTGTTGCTTCAGAGATATGGACTCCCTTACTTTTGGGGCT 600

QY 601 CCACGGTTCCTTCGTCAATTAATGATCCAAAGTTCCAGAAATACCTGTGATGAATTT 660
DB 601 CCACGGTTCCTTCGTCAATTAATGATCCAAAGTTCCAGAAATACCTGTGATGAATTT 660
QY 661 GATCTTCCCAAGGTTTGGAGGAGCTTGAACCTCACCATGACCAAGTTCATTGATTTGTC 720
DB 661 GATCTTCCCAAGGTTTGGAGGAGCTTGAACCTCACCATGACCAAGTTCATTGATTTGTC 720
QY 721 ATCCTGTGTGATGTGACTATTTGATAGCATCAAAAGTATCGGGGGCAAAAGCTCTG 780
DB 721 ATCCTGTGTGATGTGACTATTTGATAGCATCAAAAGTATCGGGGGCAAAAGCTCTG 780
QY 781 AAACCTTATTCGTCAACATGGTCCATAGAAAGCATCTTGGAGATCTTAAATAAGACAGA 840
DB 781 AAACCTTATTCGTCAACATGGTCCATAGAAAGCATCTTGGAGATCTTAAATAAGACAGA 840
QY 841 TATCAAAATTCCTGAGGACTGCGCTTACCAGAAAGCTCGACGCTTGTTCGAAGGCTAAT 900
DB 841 TATCAAAATTCCTGAGGACTGCGCTTACCAGAAAGCTCGACGCTTGTTCGAAGGCTAAT 900
QY 901 GTCACATTTGGATATTCCTGAGCTAAAATGGACTGCACCTGATGAGGAGGTCTCAAGT 960
DB 901 GTCACATTTGGATATTCCTGAGCTAAAATGGACTGCACCTGATGAGGAGGTCTCAAGT 960
QY 961 TTCCTGTGTAAGAGATAATGCTTCAACGAAGATCGGGTGACAAAGGCCATAGAGAATC 1020
DB 961 TTCCTGTGTAAGAGATAATGCTTCAACGAAGATCGGGTGACAAAGGCCATAGAGAATC 1020
QY 1021 AAATCTGCCAAGATAAATCGTCGCAAGGAGACTCGAGTCTCTTTTCAAGCCAACTGCC 1080
DB 1021 AAATCTGCCAAGATAAATCGTCGCAAGGAGACTCGAGTCTCTTTTCAAGCCAACTGCC 1080
QY 1081 ACCACATCAGCAGCGCTAAAACGGAAGGAGACTTCGGATAAAACAAAGCAAGCAGCTGC 1140
DB 1081 ACCACATCAGCAGCGCTAAAACGGAAGGAGACTTCGGATAAAACAAAGCAAGCAGCTGC 1140
QY 1141 AACCAAGAAACAAAGGCTGTGGAAGAGAAATATCTTGGATGCTTGATGATAACTA 1200
DB 1141 AACCAAGAAACAAAGGCTGTGGAAGAGAAATATCTTGGATGCTTGATGATAACTA 1200
QY 1201 CGACTAGCAAGAGCAGCGTGGCGTGATCACTTCGCTTAGATTATTAACCTCCCTGTTTA 1260
DB 1201 CGACTAGCAAGAGCAGCGTGGCGTGATCACTTCGCTTAGATTATTAACCTCCCTGTTTA 1260
QY 1261 ACTCAGAGCTTTGTTAAAGTTTCCGCCATGTTTCAAGCTGGGGTAAGTTAGTTGTTG 1320
DB 1261 ACTCAGAGCTTTGTTAAAGTTTCCGCCATGTTTCAAGCTGGGGTAAGTTAGTTGTTG 1320
QY 1321 AAGAGATTGTTGTACCAAGTAACAAAACCTTATCGCTGTTTGTGAAAAAATAAAAAA 1380
DB 1321 AAGAGATTGTTGTACCAAGTAACAAAACCTTATCGCTGTTTGTGAAAAAATAAAAAA 1380
QY 1381 A 1381
DB 1381 A 1381
RESULT 2
AAA27923
ID AAA27923 standard; cdna; 1463 BP.
XX
AC AAA27923;
XX
DT 12-SEP-2000 (first entry)
DE Maize Rad2/FEN-1 cDNA.
XX
KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
endonuclease; exonuclease; DNA repair; gene targeting; ss.
XX Zea mays.
XX
FH Key Location/Qualifiers

FT CDS 85..1224
XT /*tag= a
PN WO200036109-A1.
XX 22-JUN-2000.
XX 16-NOV-1999; 99WO-US27147.
XX 15-DEC-1998; 98US-0112332.
XX (PION-) PIONEER HI-BRED INT INC.
XX Mahajan PB;
XX WPI; 2000-452026/39.
XX P-PSDB; AAY95307.
DR Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT recombination and repair in transgenic plants, e.g. for gene targeting
PT and the production of male sterile plants -
XX Claim 1; Page 69-71; 85pp; English.
PS
XX The present sequence is that of maize cDNA coding for RAD2/FEN-1
CC (see AAY95307). The corresponding RNA was isolated from immature
CC ear tissue from 2 ears of a B73 maize line. Rad2/FEN-1 is a
CC structure specific endonuclease which under certain conditions also
CC acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to
CC produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
CC The protein is involved in the regulation of DNA repair and
CC recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.
XX
SQ Sequence 1463 BP; 466 A; 292 C; 361 G; 344 T; 0 other;
Query Watch 91.7%; Score 1267; DB 21; Length 1463;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 GCCACAGCCGCCGACGAGATGGGCATCAAGGGTTTGACGAACTGCTGGCGGACAAT 75
DB 64 GCCACAGCCGCCGACGAGATGGGCATCAAGGGTTTGACGAACTGCTGGCGGACAAT 123
QY 76 GCSCCCCAAGCGCATGAAGGACGAGAAGTTCGAGAGCTACTTCGCCGCCAATAATCGCCGTC 135
DB 124 GCSCCCCAAGCGCATGAAGGACGAGAAGTTCGAGAGCTACTTCGCCGCCAATAATCGCCGTC 183
QY 136 GAGCCACAGCATGAGCATATACCACTTCTGATTTGAGTTGGAAGGACGAGCATGGAAACT 195
DB 184 GAGCCACAGCATGAGCATATACCACTTCTGATTTGAGTTGGAAGGACGAGCATGGAACT 243
QY 196 CTCACAAATGAAGCTGGTGAAGTCACTAGTATTCATTGCAAGGAATGTTCAACCGGACAATA 255
DB 244 CTCACAAATGAAGCTGGTGAAGTCACTAGTATTCATTGCAAGGAATGTTCAACCGGACAATA 303
QY 256 AGATTACTGGAAGCGGGAATCAAGCAGTTATGTTTGGATGCAAGCCTCCTGATATG 315
DB 304 AGATTACTGGAAGCGGGAATCAAGCAGTTATGTTTGGATGCAAGCCTCCTGATATG 363
QY 316 AAGAAACAAGAGCTTGCTTAAAGATACCTCAAAAAGAGATGATGCAACCAAGATCTGACT 375
DB 364 AAGAAACAAGAGCTTGCTTAAAGATACCTCAAAAAGAGATGATGCAACCAAGATCTGACT 423

QY 376 GAGCAGTAGAGTAGAGATAAAGATCGGATTCGAAATTTGAGCAAGAGACTGTAAAG 435
DB 424 GAGGCAGTAGAGTAGAGATAAAGATCGGATTCGAAATTTGAGCAAGAGACTGTAAAG 483
QY 436 GTCACAAGGCAACACACAAGCATTTAAACGGCTATTAAAGACTTTATGGGGTTCTCTGTT 495
DB 484 GTCACAAGGCAACACACAAGCATTTAAACGGCTATTAAAGACTTTATGGGGTTCTCTGTT 543
QY 496 GTAGAGGCACCTTCTTGAAGCAGAAGCAGAATGTGCAGCCCTTTTCATATAAAGCATAAGGTG 555
DB 544 GTAGAGGCACCTTCTTGAAGCAGAAGCAGAATGTGCAGCCCTTTTCATATAAAGCATAAGGTG 603
QY 556 TTGCGCTGTTGCTTTCAGAGATATGAGACTCCCTTACTTTTGGGGCTCCACGGTTCCCTTCGT 615
DB 604 TTGCGCTGTTGCTTTCAGAGATATGAGACTCCCTTACTTTTGGGGCTCCACGGTTCCCTTCGT 663
QY 616 CATTTAATGGATCCCAAGTTTCCCAAGAAATACCTGTGATGGAATTTGATGTTGCCAAGGTT 675
DB 664 CATTTAATGGATCCCAAGTTTCCCAAGAAATACCTGTGATGGAATTTGATGTTGCCAAGGTT 723
QY 676 TTGAGAGAGCTTGAACATCACCATGGACCAGTTCATTTGATTTGTCATCCTCTGTGTGATGT 735
DB 724 TTGAGAGAGCTTGAACATCACCATGGACCAGTTCATTTGATTTGTCATCCTCTGTGTGATGT 783
QY 736 GACTATTGTAGTACATCAAGGTATCGGGGGCAACAGCTCTGAAACTTATTCCTGTCAA 795
DB 784 GACTATTGTAGTACATCAAGGTATCGGGGGCAACAGCTCTGAAACTTATTCCTGTCAA 843
QY 796 CATGGGTCCATAGAAGCATCTTTGGAGAATCTTAAATAAGACAGATATCAAAATTCCTGAG 855
DB 844 CATGGGTCCATAGAAGCATCTTTGGAGAATCTTAAATAAGACAGATATCAAAATTCCTGAG 903
QY 856 GACTGGCTTACCAGAAGCTCGACCTTGTTCAGGAGCCTAATGTCACATTTGGATATT 915
DB 904 GACTGGCTTACCAGAAGCTCGACCTTGTTCAGGAGCCTAATGTCACATTTGGATATT 963
QY 916 CCTGAGCTAAATGGACTGCACCTGATGAGGAGGCTCTCATAAGCTTTCTCTGGTAAAGAT 975
DB 964 CCTGAGCTAAATGGACTGCACCTGATGAGGAGGCTCTCATAAGCTTTCTCTGGTAAAGAT 1023
QY 976 AATGGTTTCAACGAGATCGGTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAAT 1035
DB 1024 AATGGTTTCAACGAGATCGGTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAAT 1083
QY 1036 AATGCTGCGAAGGAGACTCGAGTCTCTTTTCAAGCCAACTGCCACACATCAGCACCG 1095
DB 1084 AATGCTGCGAAGGAGACTCGAGTCTCTTTTCAAGCCAACTGCCACACATCAGCACCG 1143
QY 1096 CTAAACCGAAGGAGACTTTCGGATAAACCAAGCAGGCGCTGCGAACAAGAAACAAAG 1155
DB 1144 CTAAACCGAAGGAGACTTTCGGATAAACCAAGCAGGCGCTGCGAACAAGAAACAAAG 1203
QY 1156 GCTGGTGGAAAGAGAAATAAATTTGGATGCTTGAATGTACAACTACGACTACGAAAGCAG 1215
DB 1204 GCTGGTGGAAAGAGAAATAAATTTGGATGCTTGAATGTACAACTACGACTACGAAAGCAG 1263
QY 1216 CGGTGGCGTATCAGCTTCGCTTAGATTTTAACTCCCTGTTTAACTCAGAGCTTTGGT 1275
DB 1264 CGGTGGCGTATCAGCTTCGCTTAGATTTTAACTCCCTGTTTAACTCAGAGCTTTGGT 1323
QY 1276 AAAAGTT 1282
DB 1324 AAAAGTT 1330
RESULT 3
AAA27924
ID AAA27924 standard; cDNA; 1541 BP.
XX
AC AAA27924;
XX
DT 12-SEP-2000 (first entry)
XX

DE Maize Rad2/FEN-1 cDNA.
XX Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
KW endonuclease; exonuclease; DNA repair; gene targeting; ss.
XX Zea mays.
OS
XX
XX Key Location/Qualifiers
FH 79..1218
FT CDS /*tag= a
FT
XX
XX
PN W0200036109-A1.
PD
XX 22-JUN-2000.
XX
XX 16-NOV-1999; 99WO-US27147.
XX
XX 15-DEC-1998; 98US-0112332.
PR
XX (PION-) PIONEER HI-BRED INT INC.
PA
XX Mahajan PB;
XX
XX WPI: 2000-452026/39.
DR P-PSDB; AAY95308.
DR
XX
XX Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
PT recombination and repair in transgenic plants, e.g. for gene targeting
PT and the production of male sterile plants -
XX
XX Example 1; Page 73-74; 85pp; English.
PS
XX
XX The present sequence is that of maize cDNA (ATCC PTA-533) coding
CC for RAD2/FEN-1 (see AAY95308). The corresponding RNA was isolated
CC from a B73 line seedling after a 10 day drought, heat shocked for
CC 10 hr, and allowed to recover under normal conditions. Rad2/FEN-1
CC is a structure specific endonuclease which under certain conditions
CC also acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used
CC to produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
CC The protein is involved in the regulation of DNA repair and
CC recombination in plant systems and therefore may be used for
CC improving gene targeting during further recombinant DNA protocols
CC involving plants. RAD2/FEN-1 endonucleolytic activity is essential
CC in DNA replication and nucleotide excision and repair reactions.
CC The exolytic activity is involved in double strand break repair and
CC end joining. The protein is also useful in strand exchange
CC reactions during homologous recombination. These functions may be
CC useful in gene targeting and in the production of male sterile
CC plants. The efficacy of gene targeting can be improved by the
CC overexpression of exogenous Rad2/FEN-1 while male sterile plants can
CC be produced by the down regulation of Rad2/FEN-1 expression.
XX
XX Sequence 1541 BP; 473 A; 308 C; 377 G; 383 T; 0 other;
SQ

Query Match 88.1%; Score 1216; DB 21; Length 1541;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1266; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 16 GCCACAGCCCGCAGACGATGGGATCAAGGTTTGAGGAACCTGCTGGCGGACAAT 75
DB 58 GCCACAGCCCGCAGACGATGGGATCAAGGTTTGAGGAACCTGCTGGCGGACAAT 117
QY 76 GCGCCCAAGGGGATGAAGGACAGAAAGTTTCGAGAGCTACTTCGCCCGCAAAATCGCGTC 135
DB 118 GCGCCCAAGGGGATGAAGGACAGAAAGTTTCGAGAGCTACTTCGCCCGCAAAATCGCGTC 177
QY 136 GAGCGCAGCATGAGCATATACAGTTCCTGATTGTAGTTGGAAGGACAGGATGGAACCT 195
DB 178 GAGCGCAGCATGAGCATATACAGTTCCTGATTGTAGTTGGAAGGACAGGATGGAACCT 237
QY 196 CTCACAAATGAAGCTGGTGAAGTCACTAGTTCATTTGCAAGGAATGTTCAACCGGACAATA 255
DB 238 CTCACAAATGAAGCTGGTGAAGTCACTAGTTCATTTGCAAGGAATGTTCAACCGGACAATA 297

QY 256 AGATTACTGGAAGCGGAATCAAGCCAGTTTATGTTTGTGATGCAAGGCTCCTGATATG 315
DB 298 AGATTACTGGAAGCGGAATCAAGCCAGTTTATGTTTGTGATGCAAGGCTCCTGATATG 357
QY 316 AAGAAACAAGAGCTTGTCTAAAGATACCTCAAAAGAGATGATCAACCAAGATCTGACT 375
DB 358 AAGAAACAAGAGCTTGTCTAAAGATACCTCAAAAGAGATGATCAACCAAGATCTGACT 417
QY 376 GAGCAGTAGAGGTAGGAGATAAAGATGCCATTGAAAAATTGAGCAAGAGACTGTAAAG 435
DB 418 GAGCAGTAGAGGTAGGAGATAAAGATGCCATTGAAAAATTGAGCAAGAGACTGTAAAG 477
QY 436 GTCACAAGGCAACACAAAGAGATTGTAACGGCTATTAAAGACTTATGGGGGTTCCTGTT 495
DB 478 GTCACAAGGCAACACAAAGAGATTGTAACGGCTATTAAAGACTTATGGGGGTTCCTGTT 537
QY 496 GTAGAGGCACCTTCTGAAGCAGAAGAGATGTGCAGCCCTTTGCATAAAGAGAGGTG 555
DB 538 GTAGAGGCACCTTCTGAAGCAGAAGAGATGTGCAGCCCTTTGCATAAAGAGAGGTG 597
QY 556 TTCGCTGTGCTTCAGAGATATGACTCCCTTACTTTTGGGGCTCCACGGTTCCTTCGT 615
DB 598 TTCGCTGTGCTTCAGAGATATGACTCCCTTACTTTTGGGGCTCCACGGTTCCTTCGT 657
QY 616 CATTTAATGGATCCAAAGTTCCAAAGAAATACCTGTGATGGAATTTGATGTCGAAGGTT 675
DB 658 CATTTAATGGATCCAAAGTTCCAAAGAAATACCTGTGATGGAATTTGATGTCGAAGGTT 717
QY 676 TTGGAGAGCTTGAACCTCACCATGGACAGTTCATTGATTTGTGCATCCTGTGTGGATGT 735
DB 718 TTGGAGAGCTTGAACCTCACCATGGACAGTTCATTGATTTGTGCATCCTGTGTGGATGT 777
QY 736 GACTATTGTGATAGCATCAAGGTTATCGGGGGCAAGAGCTCTGAAACTTATTTCGCAA 795
DB 778 GACTATTGTGATAGCATCAAGGTTATCGGGGGCAAGAGCTCTGAAACTTATTTCGCAA 837
QY 796 CATGGGTCATAGAAAGCATCTTGGAGAATCTTAAATAAGACAGATATCAAAATTCCTGAG 855
DB 838 CATGGGTCATAGAAAGCATCTTGGAGAATCTTAAATAAGACAGATATCAAAATTCCTGAG 897
QY 856 GACTGCCTTACCAAGAAGCTCGAGCTTGTTCGAAGGAGCTTAATGTCAATTTGATATT 915
DB 898 GACTGCCTTACCAAGAAGCTCGAGCTTGTTCGAAGGAGCTTAATGTCAATTTGATATT 957
QY 916 CCTGAGCTAAAATGGACTGCACCTGATGAGGAGGCTCTCATAAGTTTCTGGTAAAGAT 975
DB 958 CCTGAGCTAAAATGGACTGCACCTGATGAGGAGGCTCTCATAAGTTTCTGGTAAAGAT 1017
QY 976 AATGGTTTCAACGAGATCGGGTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAT 1035
DB 1018 AATGGTTTCAACGAGATCGGGTGACAAAGGCCATAGAGAAGATCAAAATCTGCCAAGAT 1077
QY 1036 AATCTGTCGAAGGAGACTCGAGTCCCTTTTCAAGGCCAACTGCCACCACATCAGCACCG 1095
DB 1078 AATCTGTCGAAGGAGACTCGAGTCCCTTTTCAAGGCCAACTGCCACCACATCAGCACCG 1137
QY 1096 CTAAAAACGGAAGGAGACTTCGGATAAAACAAGCAGGAGCTGCCAACAAGAAACAAG 1155
DB 1138 CTAAAAACGGAAGGAGACTTCGGATAAAACAAGCAGGAGCTGCCAACAAGAAACAAG 1197
QY 1156 GCTGGTGGAAAGAAATAAATCTTGATGCTTGTATGTACAACTACGACTACGAAAGCAG 1215
DB 1198 GCTGGTGGAAAGAAATAAATCTTGATGCTTGTATGTACAACTACGACTACGAAAGCAG 1257
QY 1216 CGTGCGGTGATCAGTTCGCTTATGATTTAACTCCCTGTTTAACTCAGAGCTTTGGT 1275
DB 1258 CGTGCGGTGATCAGTTCGCTTATGATTTAACTCCCTGTTTAACTCAGAGCTTTGGT 1317
QY 1276 AAAAGTT 1282
DB 1318 AAAAGTT 1324

```

RESULT 4
AA27926
ID AAA27926 standard; cDNA; 1478 BP.
XX
AC AAA27926;
XX
DT 12-SEP-2000 (first entry)
XX
DE Maize Rad2/FEN-1 cDNA.
XX
KW Maize; Rad2/FEN-1; transgenic plant; male sterile plant;
KW endonuclease; exonuclease; DNA repair; gene targeting; ss.
XX
OS Zea mays.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 97..1236
FT /*tag= a
XX
XX WO200036109-A1.
XX
XX 22-JUN-2000.
XX
XX 16-NOV-1999; 99WO-US27147.
XX
XX 15-DEC-1998; 98US-0112332.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Mahajan PB;
XX
XX WPI; 2000-452026/39.
XX
XX P-PSDB; AAY95310.
XX
XX Maize RAD2/FEN-1 nucleic acids and proteins useful for modulating DNA
XX recombination and repair in transgenic plants, e.g. for gene targeting
XX and the production of male sterile plants -
XX
XX Example 1; Page 77-79; 85pp; English.
XX
XX The present sequence is that of maize cDNA coding for RAD2/FEN-1
XX (see AAY95310). The cDNA was isolated from a library prepared
XX from B73 line seed vitreous endosperm RNA. Rad2/FEN-1 is a
XX structure specific endonuclease which under certain conditions also
XX acts as an exonuclease. Rad2/FEN-1 nucleic acids can be used to
XX produce the Rad2/FEN-1 polypeptides in transgenic plant cells.
XX The protein is involved in the regulation of DNA repair and
XX recombination in plant systems and therefore may be used for
XX improving gene targeting during further recombinant DNA protocols
XX involving plants. RAD2/FEN-1 endonucleolytic activity is essential
XX in DNA replication and nucleotide excision and repair reactions.
XX The exolytic activity is involved in double strand break repair and
XX end joining. The protein is also useful in strand exchange
XX reactions during homologous recombination. These functions may be
XX useful in gene targeting and in the production of male sterile
XX plants. The efficacy of gene targeting can be improved by the
XX overexpression of exogenous Rad2/FEN-1 while male sterile plants can
XX be produced by the down regulation of Rad2/FEN-1 expression.
XX
XX SQ Sequence 1478 BP; 463 A; 302 C; 365 G; 348 T; 0 other;
XX
Query Match 61.5%; Score 850; DB 21; Length 1478;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1200; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX 16 GCCACAGCCCGCAGACGAGATGGCATCAAGGTTTGACGAACTGCTGGCGACAAT 75
XX |
XX 76 GCCACAGCCCGCAGACGAGATGGCATCAAGGTTTGACGAACTGCTGGCGACAAT 135
XX |
XX 76 GCCCCCAAGCGATGAGGAGAGAGTTCGAGAGCTACTTCGCCGCCAANAATCGCGCTC 135
XX |
XX 136 GCCCCCAAGCGATGAGGAGAGAGTTCGAGAGCTACTTCGCCGCCAANAATCGCGCTC 195

```

```

QY 1216 CGGTGGC 1222
Db 1276 CGGTGGC 1282

RESULT 5
AA23270
ID AAX23270 standard; DNA; 1096 BP.
AC AAX23270;
XX
XX 11-JUN-1999 (first entry)
XX
XX A. niger 2,3-dihydroxybenzoic acid decarboxylase DNA.
XX
XX 2,3-dihydroxybenzoic acid decarboxylase; aromatic hydrocarbon;
KW catechol; pharmaceutical; L-DOPA; adrenaline; agrobiological;
KW carbofuran; antioxidant; 4-tert-butyl catechol; veratol; ss.
XX
XX Aspergillus niger.
OS
XX
XX Key Location/Qualifiers
FH 1.1029
FT CDS /*tag= a
FT /product= "2,3-dihydroxybenzoic acid decarboxylase"
FT
XX
XX WO9909048-A1.
XX
XX 25-FEB-1999.
XX
XX 19-AUG-1998; 98WO-US17187.
XX
XX 18-AUG-1998; 98US-0056621.
XX
XX 20-AUG-1997; 97US-0056621.
XX
XX (UYNE-) UNIV NEBRASKA.
XX
XX Ramakrishnan S;
XX
XX WPI; 1999-190149/16.
XX
XX P-PSDB; AAW93483.
XX
XX New gene encoding 2,3-dihydroxybenzoic acid decarboxylase - useful to
PT remove carboxyl groups from ring mounted carboxyls, particularly in
PT the production of catechol
XX
XX Claim 2; Page 23-25; 30pp; English.
XX
XX This invention describes a novel Aspergillus niger 2,3-dihydroxybenzoic
XX acid decarboxylase. The products of the invention can be used in a
XX method to produce aromatic hydrocarbons, particularly catechol which
XX is used in the development of pharmaceuticals such as L-DOPA and
XX adrenaline, agrobiologicals such as carbofuran, and antioxidants
XX such as 4-tert-butyl catechol and veratol. Unlike the invention,
XX prior art aromatic hydrocarbon production methods involve the use of
XX non-renewable resources, have a negative impact on the environment, and
XX involve high temperatures and acid conditions which pose environmental
XX and safety concerns, are expensive and are energy consuming.
XX
XX Sequence 1096 BP; 264 A; 287 C; 297 G; 248 T; 0 other;
SQ
Query Match 2.0%; Score 27; DB 20; Length 1096;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1355 CTGTTTTTGAAGAAAAAAGAAAAA 1381
Db 1067 CTGTTTTTGAAGAAAAAAGAAAAA 1093

RESULT 6
AA185371
ID AAI85371 standard; cDNA; 392 BP.

```

```

XX
AC AAI85371;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 5431.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
XX
XX 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang Yt, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX P-PSDB; AAO05440.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
XX Claim 1; SEQ ID NO 5431; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 392 BP; 136 A; 71 C; 71 G; 114 T; 0 other;
SQ
Query Match 1.8%; Score 25; DB 22; Length 392;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTGTGAAAAAAGAAAAAAGAAAAA 1381
Db 88 GTTTTGTGAAAAAAGAAAAAAGAAAAA 112

RESULT 7
AAC74360
ID AAC74360 standard; cDNA; 1255 BP.
XX
XX AAC74360;
AC
XX
XX 02-FEB-2001 (first entry)
XX
XX Human secreted protein gene 24 SEQ ID NO:34.
DE
XX
XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW

```

KW nontropic; neuroprotective; antiviral; anti-allergic; hepatotropic;
 KW antidiabetic; anti-inflammatory; anti-ulcer; vulnary; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
 KW food additive; preservative; chromosome identification; cancer;
 KW female reproductive system disorder; immune disorder; wound healing;
 KW cardiovascular disorder; neurological disease; infectious disease;
 KW infection; ss.
 XX
 OS Homo sapiens.
 PN WO200058340-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 23-MAR-2000; 2000WO-US07724.
 XX
 PR 26-MAR-1999; 99US-0126510.
 PR 07-JAN-2000; 2000US-0174850.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Komatsoulis G;
 XX
 DR WPI; 2000-594638/56.
 DR P-PSDB; AAB39425.
 XX
 PF Fifty nucleic acid molecules encoding human secreted proteins, useful
 PT in the prevention, treatment and diagnosis of cancer, immune disorders,
 PT cardiovascular disorders and neurological diseases -
 XX
 PS Claim 1; Page 341-342; 391pp; English.
 CC The polynucleotide sequences given in AAC74337 to AAC74386 encode the
 CC human secreted proteins given in AAB39402 to AAB39451. AAB39452 to
 CC AAB39484 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Example of activities include:
 CC cytostatic; immunosuppressive; nontropic; neuroprotective; antiviral;
 CC anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
 CC vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
 CC cardiant. The polynucleotides and polypeptides are useful for preventing,
 CC treating or ameliorating a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. The polypeptides
 CC can also be used as a food additive or preservative to increase or
 CC decrease storage capabilities. The polynucleotide are useful for
 CC chromosome identification. They are also useful as probes for diagnosing
 CC a disorder related to the female reproductive system, particularly breast
 CC and/or ovarian cancer. They are also useful in the gene therapy of breast
 CC and ovarian cancer. Secreted protein nucleic acids, proteins,
 CC antibodies, agonists and antagonists are useful in the diagnosis,
 CC treatment and prevention of: (a) cancer; (b) immune disorders; (c)
 CC cardiovascular disorders; (d) wound healing; (e) neurological diseases;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections. AAC74328 to AAC74336 and AAB39401 represent
 CC sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 1255 BP; 357 A; 232 C; 273 G; 393 T; 0 other;
 Query Match 1.8%; Score 25; DB 21; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1357 GTTTTTCAGAAAAAAGAAAAA 1381
 |||||||
 DB 1219 GTTTTTCAGAAAAAAGAAAAA 1243
 |||||||
 RESULT 8
 AAC76262
 ID AAC76262 standard; cDNA; 1570 BP.
 XX
 AC AAC76262;

XX
 DT 08-FEB-2001 (first entry)
 XX
 DE Human ORFX ORF1917 polynucleotide sequence SEQ ID NO:3633.
 XX
 KW Human; open reading frame: ORFX; detection; cytostatic; hepatotropic;
 KW vulnary; antiparasitic; antiparkinsonian; nontropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; anti-inflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; anti-inflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000WO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR P-PSDB; AAB42053.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 5; Page 2791-2792; 5507pp; English.
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiparasitic; antiparkinsonian; nontropic; neuroprotective;
 CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC anti-inflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antithyroid; and antianaemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorder, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, anti-inflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 1570 BP; 401 A; 398 C; 440 G; 328 T; 3 other;

Query Match 1.8%; Score 25; DB 21; Length 1570;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1357 GTTTTTCAGAAAAA 1381
1358 TTTTTCAGAAAAA 1382
1359 TTTTTCAGAAAAA 1383
1360 TTTTTCAGAAAAA 1384
1361 TTTTTCAGAAAAA 1385
1362 TTTTTCAGAAAAA 1386
1363 TTTTTCAGAAAAA 1387
1364 TTTTTCAGAAAAA 1388
1365 TTTTTCAGAAAAA 1389
1366 TTTTTCAGAAAAA 1390
1367 TTTTTCAGAAAAA 1391
1368 TTTTTCAGAAAAA 1392
1369 TTTTTCAGAAAAA 1393
1370 TTTTTCAGAAAAA 1394
1371 TTTTTCAGAAAAA 1395
1372 TTTTTCAGAAAAA 1396
1373 TTTTTCAGAAAAA 1397
1374 TTTTTCAGAAAAA 1398
1375 TTTTTCAGAAAAA 1399
1376 TTTTTCAGAAAAA 1400
1377 TTTTTCAGAAAAA 1401
1378 TTTTTCAGAAAAA 1402
1379 TTTTTCAGAAAAA 1403
1380 TTTTTCAGAAAAA 1404
1381 TTTTTCAGAAAAA 1405
1382 TTTTTCAGAAAAA 1406
1383 TTTTTCAGAAAAA 1407
1384 TTTTTCAGAAAAA 1408
1385 TTTTTCAGAAAAA 1409
1386 TTTTTCAGAAAAA 1410
1387 TTTTTCAGAAAAA 1411
1388 TTTTTCAGAAAAA 1412
1389 TTTTTCAGAAAAA 1413
1390 TTTTTCAGAAAAA 1414
1391 TTTTTCAGAAAAA 1415
1392 TTTTTCAGAAAAA 1416
1393 TTTTTCAGAAAAA 1417
1394 TTTTTCAGAAAAA 1418
1395 TTTTTCAGAAAAA 1419
1396 TTTTTCAGAAAAA 1420
1397 TTTTTCAGAAAAA 1421
1398 TTTTTCAGAAAAA 1422
1399 TTTTTCAGAAAAA 1423
1400 TTTTTCAGAAAAA 1424
1401 TTTTTCAGAAAAA 1425
1402 TTTTTCAGAAAAA 1426
1403 TTTTTCAGAAAAA 1427
1404 TTTTTCAGAAAAA 1428
1405 TTTTTCAGAAAAA 1429
1406 TTTTTCAGAAAAA 1430
1407 TTTTTCAGAAAAA 1431
1408 TTTTTCAGAAAAA 1432
1409 TTTTTCAGAAAAA 1433
1410 TTTTTCAGAAAAA 1434
1411 TTTTTCAGAAAAA 1435
1412 TTTTTCAGAAAAA 1436
1413 TTTTTCAGAAAAA 1437
1414 TTTTTCAGAAAAA 1438
1415 TTTTTCAGAAAAA 1439
1416 TTTTTCAGAAAAA 1440
1417 TTTTTCAGAAAAA 1441
1418 TTTTTCAGAAAAA 1442
1419 TTTTTCAGAAAAA 1443
1420 TTTTTCAGAAAAA 1444
1421 TTTTTCAGAAAAA 1445
1422 TTTTTCAGAAAAA 1446
1423 TTTTTCAGAAAAA 1447
1424 TTTTTCAGAAAAA 1448
1425 TTTTTCAGAAAAA 1449
1426 TTTTTCAGAAAAA 1450
1427 TTTTTCAGAAAAA 1451
1428 TTTTTCAGAAAAA 1452
1429 TTTTTCAGAAAAA 1453
1430 TTTTTCAGAAAAA 1454
1431 TTTTTCAGAAAAA 1455
1432 TTTTTCAGAAAAA 1456
1433 TTTTTCAGAAAAA 1457
1434 TTTTTCAGAAAAA 1458
1435 TTTTTCAGAAAAA 1459
1436 TTTTTCAGAAAAA 1460
1437 TTTTTCAGAAAAA 1461
1438 TTTTTCAGAAAAA 1462
1439 TTTTTCAGAAAAA 1463
1440 TTTTTCAGAAAAA 1464
1441 TTTTTCAGAAAAA 1465
1442 TTTTTCAGAAAAA 1466
1443 TTTTTCAGAAAAA 1467
1444 TTTTTCAGAAAAA 1468
1445 TTTTTCAGAAAAA 1469
1446 TTTTTCAGAAAAA 1470
1447 TTTTTCAGAAAAA 1471
1448 TTTTTCAGAAAAA 1472
1449 TTTTTCAGAAAAA 1473
1450 TTTTTCAGAAAAA 1474
1451 TTTTTCAGAAAAA 1475
1452 TTTTTCAGAAAAA 1476
1453 TTTTTCAGAAAAA 1477
1454 TTTTTCAGAAAAA 1478
1455 TTTTTCAGAAAAA 1479
1456 TTTTTCAGAAAAA 1480
1457 TTTTTCAGAAAAA 1481
1458 TTTTTCAGAAAAA 1482
1459 TTTTTCAGAAAAA 1483
1460 TTTTTCAGAAAAA 1484
1461 TTTTTCAGAAAAA 1485
1462 TTTTTCAGAAAAA 1486
1463 TTTTTCAGAAAAA 1487
1464 TTTTTCAGAAAAA 1488
1465 TTTTTCAGAAAAA 1489
1466 TTTTTCAGAAAAA 1490
1467 TTTTTCAGAAAAA 1491
1468 TTTTTCAGAAAAA 1492
1469 TTTTTCAGAAAAA 1493
1470 TTTTTCAGAAAAA 1494
1471 TTTTTCAGAAAAA 1495
1472 TTTTTCAGAAAAA 1496
1473 TTTTTCAGAAAAA 1497
1474 TTTTTCAGAAAAA 1498
1475 TTTTTCAGAAAAA 1499
1476 TTTTTCAGAAAAA 1500
1477 TTTTTCAGAAAAA 1501
1478 TTTTTCAGAAAAA 1502
1479 TTTTTCAGAAAAA 1503
1480 TTTTTCAGAAAAA 1504
1481 TTTTTCAGAAAAA 1505
1482 TTTTTCAGAAAAA 1506
1483 TTTTTCAGAAAAA 1507
1484 TTTTTCAGAAAAA 1508
1485 TTTTTCAGAAAAA 1509
1486 TTTTTCAGAAAAA 1510
1487 TTTTTCAGAAAAA 1511
1488 TTTTTCAGAAAAA 1512
1489 TTTTTCAGAAAAA 1513
1490 TTTTTCAGAAAAA 1514
1491 TTTTTCAGAAAAA 1515
1492 TTTTTCAGAAAAA 1516
1493 TTTTTCAGAAAAA 1517
1494 TTTTTCAGAAAAA 1518
1495 TTTTTCAGAAAAA 1519
1496 TTTTTCAGAAAAA 1520
1497 TTTTTCAGAAAAA 1521
1498 TTTTTCAGAAAAA 1522
1499 TTTTTCAGAAAAA 1523
1500 TTTTTCAGAAAAA 1524
1501 TTTTTCAGAAAAA 1525
1502 TTTTTCAGAAAAA 1526
1503 TTTTTCAGAAAAA 1527
1504 TTTTTCAGAAAAA 1528
1505 TTTTTCAGAAAAA 1529
1506 TTTTTCAGAAAAA 1530
1507 TTTTTCAGAAAAA 1531
1508 TTTTTCAGAAAAA 1532
1509 TTTTTCAGAAAAA 1533
1510 TTTTTCAGAAAAA 1534
1511 TTTTTCAGAAAAA 1535
1512 TTTTTCAGAAAAA 1536
1513 TTTTTCAGAAAAA 1537
1514 TTTTTCAGAAAAA 1538
1515 TTTTTCAGAAAAA 1539
1516 TTTTTCAGAAAAA 1540
1517 TTTTTCAGAAAAA 1541
1518 TTTTTCAGAAAAA 1542
1519 TTTTTCAGAAAAA 1543
1520 TTTTTCAGAAAAA 1544
1521 TTTTTCAGAAAAA 1545
1522 TTTTTCAGAAAAA 1546
1523 TTTTTCAGAAAAA 1547
1524 TTTTTCAGAAAAA 1548
1525 TTTTTCAGAAAAA 1549
1526 TTTTTCAGAAAAA 1550
1527 TTTTTCAGAAAAA 1551
1528 TTTTTCAGAAAAA 1552
1529 TTTTTCAGAAAAA 1553
1530 TTTTTCAGAAAAA 1554
1531 TTTTTCAGAAAAA 1555
1532 GTTTTTCAGAAAAA 1556

RESULT 9
AAC59600
ID AAC59600 standard; cDNA; 1648 BP.
XX
XX AAC59600;
XX 26-JAN-2001 (first entry)
XX Human secreted protein gene 35 SEQ ID NO:45.
XX
XX Human; secreted protein; diagnosis; neuroprotective; cytostatic;
KW cardioactive; immunomodulatory; muscular active general; vulnerary;
KW gastrointestinal; nephrotropic; antineoplastic; gynaecological; and
KW antibacterial; gene therapy; detection; cancer; chromosome marker;
KW chromosome identification; neural disorder; immune disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW proliferative disorder; wound healing; infectious disease; preservative;
KW food additive; ss.
XX
XX Homo sapiens.
XX WO2000056883-A1.
XX
XX 28-SEP-2000.
XX
XX 16-MAR-2000; 2000WO-US06822.
XX
XX 23-MAR-1999; 99US-0126054.
XX 10-DEC-1999; 99US-0169916.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis G;
XX WPI; 2000-587666/55.
XX P-PSDB; AAB34333.
XX
XX Human secreted proteins and gene sequences encoding them, useful for
XX detecting, preventing, and treating disorders such as cancer,
XX neurological disorders and immune system disorders -
XX
XX Claim 1; Page 359-360; 429pp; English.
XX
XX The polynucleotide sequences given in AAC59566 to AAC59614 encode the
XX human secreted proteins given in AAB34299 to AAB34347. AAB34348 to
XX AAB34437 represent human secreted polypeptide sequences and proteins
XX homologous to them, which are given in the exemplification of the present
XX invention. Human secreted proteins have activities based on the tissues
XX and cells the genes are expressed in. Example of activities include:
XX neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular
XX active general; vulnerary; gastrointestinal; nephrotropic;
XX antineoplastic; gynaecological; and antibacterial. The polynucleotides
XX can be used for the detection of various disorders such as cancer,
XX chromosome identification, as chromosome markers, and for numerous other
XX diagnostic or research purposes. The secreted proteins can be used to
XX treat disorders such as neural, immune, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX disorders, wound healing, and infectious diseases. The proteins can
XX also be used as a food additive or preservative to increase or decrease
XX storage capabilities. AAC59557 to AAC59565 and AAB34298 represent
XX sequences used in the exemplification of the present invention.
XX
XX Sequence 1648 BP; 408 A; 419 C; 467 G; 354 T; 0 other;

CC 28 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g. radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein-encoding cDNA of the invention.

XX Sequence 343 BP; 143 A; 50 C; 68 G; 82 T; 0 other;

Query Match 1.7%; Score 24; DB 22; Length 343;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAA 1381
 Db 258 TTTTGTGAAAAA 281

RESULT 11
 AAF64376/c
 ID AAF64376 standard; cDNA: 387 BP.

XX AAF64376;

XX 09-APR-2001 (first entry)

XX Novel human polynucleotide, SEQ ID NO: 132.

XX Human; cytostatic; gene therapy; colon cancer; prostate cancer;
 KW breast cancer; lung cancer; cancer detection; ss.

XX Homo sapiens.

XX WO200102568-A2.

XX 11-JAN-2001.

XX 30-JUN-2000; 2000WO-US18374.

XX 02-JUL-1999; 99US-0142310.

XX 02-JUL-1999; 99US-0142311.

XX (CHIR) CHIRON CORP.

XX (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klingner J, Kassam A;
 PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G, Drmanac R;
 PI Crkenjakov R, Drmanac S, Dickson M, Labat I, Leshkowitz D;
 PI Kita D, Garcia V, Jones LW, Strache-Crain B;

XX WPI: 2001-091805/10.

XX Library of polynucleotides for diagnosing a cancerous state of a
 PT mammalian cell and detecting cancer, particularly of the colon or
 PT prostate, comprises 3351 human polynucleotide sequences -

XX Claim 9; Page 563; 1046pp; English.

XX

CC The present sequence is one of 3351 sequences in a library of human
 CC polynucleotides. The library is used to detect differentially expressed
 CC genes correlated with a cancerous state of a mammalian cell and can
 CC detect colon, prostate, breast and lung cancer. The library can be used
 CC to produce probes for detection of mRNA and to produce additional copies
 CC of the polynucleotides. The probes can be used for chromosome mapping of
 CC the polynucleotide and for detection of transcription levels. Ribozymes
 CC or antisense oligonucleotides can be generated. The polynucleotides and
 CC their gene products are used as genetic or biochemical markers (e.g. in
 CC blood or tissues) that will detect the earliest changes along the
 CC carcinogenesis pathway and/or monitor the efficacy of therapies and
 CC preventive interventions. The polynucleotides, polypeptides and
 CC antibodies against them can be used in pharmaceutical compositions to
 CC treat the cancers and proliferative disorders such as neoplasia,
 CC dysplasia and hyperplasia.

XX Sequence 387 BP; 127 A; 63 C; 97 G; 100 T; 0 other;

Query Match 1.7%; Score 24; DB 22; Length 387;

Best Local Similarity 100.0%; Pred. No. 5.7;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAA 1381
 Db 69 TTTTGTGAAAAA 46

RESULT 12
 AA182336/c
 ID AA182336 standard; cDNA: 403 BP.

XX AA182336;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 2396.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation; ss.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US04927.

XX 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI: 2001-514838/56.

XX P-PSDB; AA002405.

XX Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukaemia, inflammation and immune
 PT disorders -

XX Claim 1; SEQ ID NO 2396; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating

activity, tissue growth factor activity, immunomodulatory activity and
activin/inhibin activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

Sequence 403 BP; 82 A; 131 C; 115 G; 74 T; 1 other;

Query Match 1.7%; Score 24; DB 22; Length 403;
Best Local Similarity 100.0%; Pred. NO. 5.7; 0; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;

QY 1358 TTTTGTGAAAAAAGAAAAAAAAAA 1381
|||||TTTTTTTTTTTTTTTTTTTT
DB 99 TTTTGTGAAAAAAGAAAAAAAAAA 76

RESULT 13
AAI89042
ID ID AAI89042 standard; cDNA; 405 BP.
XX AC AAI89042;
AC XX
XX DT
DT 06-NOV-2001 (first entry)
XX Human polynucleotide SEQ ID NO 9102.
DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation; ss.
XX Homo sapiens.
OS
XX WO200164835-A2.
PN
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
PI WPI; 2001-514838/56.
XX P-PSDB; AA009111.
PT Isolated nucleic acids and polypeptides, useful for preventing
disorders -
XX
XX Claim 1; SEQ ID NO 9102; 1399pp + Sequence Listing: English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and
activin/inhibin activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation.
Note: The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences.

```
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 28-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 12-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244517.
PR 08-NOV-2000; 2000US-0245474.
PR 08-NOV-2000; 2000US-0245475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.

PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX WPI; 2001-483426/52.
DR P-PSDB; AAM85334.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Claim 1; SEQ ID NO 3175; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
XX SQ Sequence 406 BP; 107 A; 99 C; 74 G; 123 T; 3 other;

Query Match 1.7%; Score 24; DB 22; Length 406;
Best Local Similarity 100.0%; Pred.No. 5.7;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAA 1381
IIIIIIIIIIIIIIIIIIII
DB 380 TTTTGTGAAAAA 403

RESULT 15
AAI92294
ID AAI92294 standard; cDNA; 425 BP.
XX
AC AAI92294;
XX
XX 06-NOV-2001 (first entry)
XX
```

```

DE Human polynucleotide SEQ ID NO 12354.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
XX
XX 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
XX
XX 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-514838/56.
XX
XX P-PSDB; AAO12363.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
XX Claim 1; SEQ ID NO 12354; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 425 BP; 141 A; 80 C; 124 G; 80 T; 0 other;
SQ
Query Match 1.7%; Score 24; DB 22; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.6;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1358 TTTTGTGAAAAA 1381
Db 281 TTTTGTGAAAAA 304
RESULT 16
AAK66234
ID AAK66234 standard; DNA; 469 BP.
XX
XX AAK66234;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:21046.
DE
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
OS
XX WO200157182-A2.
PN

```

```

XX 09-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US01354.
PF
XX
XX 31-JAN-2000; 2000US-0179065.
XX
XX 04-FEB-2000; 2000US-0180628.
PR
XX 24-FEB-2000; 2000US-0184664.
PR
XX 02-MAR-2000; 2000US-0186350.
PR
XX 16-MAR-2000; 2000US-0189874.
PR
XX 17-MAR-2000; 2000US-0190076.
PR
XX 18-APR-2000; 2000US-0198123.
PR
XX 19-MAY-2000; 2000US-0205515.
PR
XX 07-JUN-2000; 2000US-0209467.
PR
XX 28-JUN-2000; 2000US-0214886.
PR
XX 30-JUN-2000; 2000US-0215135.
PR
XX 07-JUL-2000; 2000US-0216647.
PR
XX 07-JUL-2000; 2000US-0216880.
PR
XX 11-JUL-2000; 2000US-0217487.
PR
XX 11-JUL-2000; 2000US-0217496.
PR
XX 14-JUL-2000; 2000US-0218290.
PR
XX 26-JUL-2000; 2000US-0220963.
PR
XX 26-JUL-2000; 2000US-0220964.
PR
XX 14-AUG-2000; 2000US-0224518.
PR
XX 14-AUG-2000; 2000US-0224519.
PR
XX 14-AUG-2000; 2000US-0225213.
PR
XX 14-AUG-2000; 2000US-0225214.
PR
XX 14-AUG-2000; 2000US-0225266.
PR
XX 14-AUG-2000; 2000US-0225267.
PR
XX 14-AUG-2000; 2000US-0225268.
PR
XX 14-AUG-2000; 2000US-0225270.
PR
XX 14-AUG-2000; 2000US-0225447.
PR
XX 14-AUG-2000; 2000US-0225757.
PR
XX 14-AUG-2000; 2000US-0225758.
PR
XX 14-AUG-2000; 2000US-0225759.
PR
XX 18-AUG-2000; 2000US-0226279.
PR
XX 22-AUG-2000; 2000US-0226681.
PR
XX 22-AUG-2000; 2000US-0226888.
PR
XX 22-AUG-2000; 2000US-0227182.
PR
XX 23-AUG-2000; 2000US-0227009.
PR
XX 30-AUG-2000; 2000US-0228924.
PR
XX 01-SEP-2000; 2000US-0229287.
PR
XX 01-SEP-2000; 2000US-0229343.
PR
XX 01-SEP-2000; 2000US-0229344.
PR
XX 01-SEP-2000; 2000US-0229345.
PR
XX 05-SEP-2000; 2000US-0229509.
PR
XX 05-SEP-2000; 2000US-0229513.
PR
XX 06-SEP-2000; 2000US-0230437.
PR
XX 06-SEP-2000; 2000US-0230438.
PR
XX 08-SEP-2000; 2000US-0231242.
PR
XX 08-SEP-2000; 2000US-0231243.
PR
XX 08-SEP-2000; 2000US-0231244.
PR
XX 08-SEP-2000; 2000US-0231413.
PR
XX 08-SEP-2000; 2000US-0231414.
PR
XX 08-SEP-2000; 2000US-0232080.
PR
XX 08-SEP-2000; 2000US-0232081.
PR
XX 12-SEP-2000; 2000US-0231968.
PR
XX 14-SEP-2000; 2000US-0232397.
PR
XX 14-SEP-2000; 2000US-0232398.
PR
XX 14-SEP-2000; 2000US-0232399.
PR
XX 14-SEP-2000; 2000US-0232400.
PR
XX 14-SEP-2000; 2000US-0232401.
PR
XX 14-SEP-2000; 2000US-0233063.
PR
XX 14-SEP-2000; 2000US-0233064.
PR
XX 14-SEP-2000; 2000US-0233065.
PR
XX 21-SEP-2000; 2000US-0234223.
PR
XX 21-SEP-2000; 2000US-0234274.
PR
XX 25-SEP-2000; 2000US-0234997.
PR
XX 25-SEP-2000; 2000US-0234998.
PR
XX 26-SEP-2000; 2000US-0235484.
PR
XX 27-SEP-2000; 2000US-0235834.
PR
XX 27-SEP-2000; 2000US-0235836.
PR
XX 29-SEP-2000; 2000US-0236327.
PR

```


DR P-PSDB; AAB44214.

XX Novel isolated nucleic acids comprising sequences encoding peptides

PT useful for treating or diagnosing e.g. cancer -

PS Claim 1; Page 1371; 2352pp; English.

XX

XX AAC77607 to AAC78448 encode the human cancer associated proteins given

CC in AAB43398 to AAB44239. The proteins can have activities based on the

CC tissues and cells the genes are expressed in. Example of activities

CC include: cytostatic; proliferative; vulnerary; immunomodulator;

CC anti-diabetic; antiasthmatic; antirheumatic; antiarthritic;

CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;

CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;

CC nootropic; vasotropic; antipsoriatic and antiangiogenic. The

CC polynucleotides and polypeptides can be used for preventing, treating or

CC ameliorating medical conditions and diagnosing pathological conditions.

CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from

CC the present invention may be used to treat immune disorders by activating

CC or inhibiting the proliferation, differentiation or mobilisation of

CC immune cells, to treat disorders of haematopoietic cells, autoimmune

CC disorders, allergic reactions, graft versus host disease and organ

CC rejection, modulate haemostatic or thrombolytic activity, modulate

CC inflammation, cancers, cardiovascular disorders, neurological disease and

CC bacterial or viral infections. The peptides, nucleotides, antibodies,

CC agonists and antagonists may be also be used in drug screens. AAC78449 to

CC AAC78457 and AAB44240 represent sequences used in the exemplification of

CC the present invention.

XX

XX Sequence 500 BP; 111 A; 128 C; 145 G; 86 T; 30 other;

SQ

Query Match 1.7%; Score 24; DB 21; Length 500;

Best Local Similarity 100.0%; Pred. No. 5.5;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGACCCACGGTCGCGCCACAGCC 24

Db 2 CGACCCACGGTCGCGCCACAGCC 25

|||||

RESULT 18

AAH70748

ID AAH70748 standard; cDNA; 573 BP.

XX

AC AAH70748;

XX

XX 19-SEP-2001 (first entry)

DT

XX

XX Human cervical cancer marker nucleic acid 2022.

DE

XX

XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

KW

XX

XX Homo sapiens.

OS

XX

XX WO200142467-A2.

PN

XX

XX 14-JUN-2001.

PD

XX

XX 08-DEC-2000; 2000WO-US33312.

PF

XX

XX 08-DEC-1999; 99US-0169681.

PR

XX

XX 21-DEC-1999; 99US-0171350.

PR

XX

XX 14-MAR-2000; 2000US-0189315.

PR

XX

XX 12-MAY-2000; 2000US-0203791.

PR

XX

XX 09-JUN-2000; 2000US-0210600.

PR

XX

XX 21-JUL-2000; 2000US-0220114.

PR

XX

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PA

XX

XX Schlegel R, Deeds J, Berger A, Zhao X;

PI

XX

XX WPI; 2001-375006/39.

DR

XX

PT New isolated nucleic acid for diagnosing and treating cervical cancer

PT and for assessing and detecting compounds for treating the cancer -

XX Claim 1; Page 431; 1051pp; English.

PS

XX

XX The invention relates to novel genes (AAH68727-AAH73383) associated with

CC cervical cancer with cytostatic activity. The nucleic acids and encoded

CC polypeptides are useful: to assess if a patient is afflicted with

CC cervical cancer or has a pre-malignant condition; to monitor the

CC progression of cervical cancer or a premalignant condition in a patient;

CC and to select and/or assess the efficacy of a compound or therapy for

CC inhibiting cervical cancer in a patient. The nucleic acids may also be

CC useful for gene therapy.

XX

SQ Sequence 573 BP; 161 A; 121 C; 133 G; 158 T; 0 other;

Query Match 1.7%; Score 24; DB 22; Length 573;

Best Local Similarity 100.0%; Pred. No. 5.4;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTTCGAAAAAATAAAAAA 1381

Db 517 TTTTTCGAAAAAATAAAAAA 540

|||||

RESULT 19

AAD05657

ID AAD05657 standard; cDNA; 579 BP.

XX

XX AAD05657;

AC

XX

XX 17-JUL-2001 (first entry)

DT

XX

XX Human secreted protein-encoding gene 27 cDNA clone HLDQA07, SEQ ID NO:89.

DE

XX

XX Human; secreted protein; proliferative disorder; cancer; tumour;

KW foetal abnormality; developmental abnormality; haematopoietic disorder;

KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;

KW inflammation; allergy; neurological disorder; Alzheimer's disease;

KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;

KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;

KW cardiovascular disorder; angiogenic disorder; kidney disorder;

KW gastrointestinal disorder; pregnancy-related disorder;

KW endocrine disorder; infection; wound healing; vulnerary;

KW cell culture; chemotaxis; food additive; gene therapy;

KW binding partner identification; ss.

XX

XX Homo sapiens.

OS

XX

XX Location/Qualifiers

FH Key 359..577

FT CDS /*tag= a

FT /*product= "Human secreted protein"

FT /*note= "CDS does not include stop codon"

FT /*partial

FT sig_peptide 359..445

FT /*tag= b

FT mat_peptide 446..577

FT /*tag= c

FT /*product= "Human mature secreted protein"

FT

XX

XX WO200134627-A1.

PN

XX

XX 17-MAY-2001.

PD

XX

XX 08-NOV-2000; 2000WO-US30628.

PF

XX

XX 12-NOV-1999; 99US-0164744.

PR

XX

XX 30-JUN-2000; 2000US-0215140.

PR

XX

XX (HUMA-) HUMAN GENOME SCI INC.

PA

XX

XX Ruben SM, Komatsoulis GA, Baker KP, Young PE;

PI

```
XX WPI; 2001-316491/33.
DR P-PSDB; AAE01848.
XX
XX New nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -
XX
XX Claim 1; Page 481; 567pp; English.
XX
XX AAD05579-AAD05658 represent cDNAs corresponding to 28 human secreted
CC protein genes and AAE01770-AAE01849 represent the proteins they encode.
CC AAE01850-AAE01860 represent human secreted protein fragments or variants.
CC The genes and their secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 28 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g. radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein-encoding cDNA of the invention.
XX
XX Sequence 579 BP; 189 A; 98 C; 131 G; 160 T; 1 other;
SQ
Query Match 1.7%; Score 24; DB 22; Length 579;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1358 TTTTGTGAAAAAATAAAAAAAAAA 1381
DB 490 TTTTGTGAAAAAATAAAAAAAAAA 513
XX
RESULT 20
AAH71472
ID AAH71472 standard; cDNA; 597 BP.
XX
XX AAH71472;
XX
XX 19-SEP-2001 (first entry)
XX
XX Human cervical cancer marker nucleic acid 2746.
XX
XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX WO200142467-A2.
XX
XX 14-JUN-2001.
XX
XX 08-DEC-2000; 2000WO-US33312.
XX
XX 08-DEC-1999; 99US-0169681.
XX
XX 21-DEC-1999; 99US-0171350.
XX
PR
```

```
PR 14-MAR-2000; 2000US-0189315.
PR 12-MAY-2000; 2000US-0203791.
PR 09-JUN-2000; 2000US-0210600.
PR 21-JUL-2000; 2000US-0220114.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Deeds J, Berger A, Zhao X;
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer
PT and for assessing and detecting compounds for treating the cancer -
XX
XX Claim 1; Page 564; 1051pp; English.
XX
XX The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy.
XX
XX Sequence 597 BP; 259 A; 54 C; 39 G; 244 T; 1 other;
SQ
Query Match 1.7%; Score 24; DB 22; Length 597;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1358 TTTTGTGAAAAAATAAAAAAAAAA 1381
DB 310 TTTTGTGAAAAAATAAAAAAAAAA 333
XX
RESULT 21
AAK71980
ID AAK71980 standard; DNA; 860 BP.
XX
XX AAK71980;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:26792.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
XX
XX 04-FEB-2000; 2000US-0180628.
XX
XX 24-FEB-2000; 2000US-0184664.
XX
XX 02-MAR-2000; 2000US-0186350.
XX
XX 16-MAR-2000; 2000US-0189874.
XX
XX 17-MAR-2000; 2000US-0190076.
XX
XX 18-APR-2000; 2000US-0198123.
XX
XX 19-MAY-2000; 2000US-0205515.
XX
XX 07-JUN-2000; 2000US-0209467.
XX
XX 28-JUN-2000; 2000US-0214886.
XX
XX 30-JUN-2000; 2000US-0215135.
XX
XX 07-JUL-2000; 2000US-0216647.
XX
XX 07-JUL-2000; 2000US-0216880.
XX
XX 11-JUL-2000; 2000US-0217487.
XX
XX 14-JUL-2000; 2000US-0217496.
XX
XX 14-JUL-2000; 2000US-0218290.
XX
PR
```

PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX Disclosure; SEQ ID NO 26792; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially

CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.

XX
SQ Sequence 860 BP; 213 A; 226 C; 219 G; 202 T; 0 other;

Query Match 1.7%; Score 24; DB 22; Length 860;

Best Local Similarity 100.0%; Pred. No. 5.1;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1358 TTTTGTGAAAAAATAAAAAA 1381

|||||

Db 550 TTTTGTGAAAAAATAAAAAA 573

RESULT 22

AAK71981

ID AAK71981 standard; DNA; 860 BP.

XX AAK71981;

AC AAK71981;

XX 06-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:26793.

DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cystostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226271.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226688.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 21-SEP-2000; 2000US-0234977.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-024617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.

PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
DR
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX Disclosure; SEQ ID NO 26793; 3071pp + Sequence Listing; English.
PS
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 860 BP; 213 A; 226 C; 219 G; 202 T; 0 other;

Query Match 1.7%; Score 24; DB 22; Length 860;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAAAGAAAAA 1381
Db 550 TTTTGTGAAAAAAGAAAAA 573

RESULT 23
AAF09266/C
ID AAF09266 standard; cDNA; 1061 BP.
XX

AC AAF09266;
XX
XX 13-MAR-2001 (first entry)
XX
XX Fusarium venenatum EST SEQ ID NO:1789.
DE
XX
XX Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
XX Fusarium venenatum.
OS
XX WO2000056762-A2.
XX
XX 28-SEP-2000.
PD
XX
XX 22-MAR-2000; 2000WO-US07781.
PF
XX
XX 22-MAR-1999; 99US-0273623.
PR
XX (NOVO) NOVO NORDISK BIOTECH INC.
PA (NOVO) NOVO NORDISK AS.
PA
XX Berka RM, Rey MW, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
PI WPI; 2000-594572/56.
XX
XX Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags -
XX
XX Claim 86; Page 1053; 3161pp; English.
PS
XX
XX The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring
CC the global expression of genes from FF cells allows the production
CC potential of the microorganisms to be improved. New genes may be
CC discovered, possible functions of unknown open reading frames can be
CC identified and gene copy number variation and stability can be
CC monitored. The expression of genes can be used to study how FF cells
CC adapt to changes in culture conditions, environmental stress, spore
CC morphogenesis, recombination, metabolic or catabolic pathway
CC engineering. Using ESTs provides several advantages over genomic or
CC random cDNA clones including elimination of redundancy as one spot on an
CC array equals one gene or open reading frame, and organisation of the
CC microarrays based on function of the gene products to facilitate
CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
CC Fusarium venenatum; AAF11248 to AAF11853 represents ESTs from Aspergillus
CC niger; AAF11854 to AAF14878 represents ESTs from Aspergillus oryzae; and
CC AAF14879 to AAF15337 represents ESTs from Trichoderma reesei, which are
CC all specifically claimed in the present invention.
XX
SQ Sequence 1061 BP; 266 A; 228 C; 303 G; 263 T; 1 other;

Query Match 1.7%; Score 24; DB 21; Length 1061;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAAAGAAAAA 1381
Db 51 TTTTGTGAAAAAAGAAAAA 28

RESULT 24
AAF09266
AAF09266

```

ID AAV30555 standard; DNA; 1304 BP.
XX
AC AAV30555;
XX
XX 14-OCT-1998 (first entry)
XX
DE Strawberry chalcone synthesis enzyme gene sequence.
XX
KW Chalcone synthesis; enzyme; strawberry; Fragaria ananassa; anthocyanin;
KW transgenic plant; colour; ds.
XX
XX Fragaria x ananassa.
OS
XX JP09308487-A.
XX
XX 02-DEC-1997.
XX
XX 22-MAY-1996; 96JP-0127245.
XX
XX 22-MAY-1996; 96JP-0127245.
XX
XX (ISHI ) ISHIKAWAJIMA HARIMA HEAVY IND.
XX
XX WPI; 1998-070648/07.
XX
XX Chalcone synthetic enzyme gene - used to produce plants of high
XX anthocyanin productivity
XX
XX Claim 1; Page 2; 5pp; Japanese.
XX
XX This sequence represents the gene encoding a chalcone synthetic enzyme
XX from strawberries (Fragaria x ananassa). The gene is used to produce
XX plant cells with increased anthocyanin productivity. This can lead to
XX the production of transformed plants which differ in colour from
XX conventional plants of that species.
XX
XX Sequence 1304 BP; 330 A; 309 C; 330 G; 335 T; 0 other;
XX
Query Match 1.7%; Score 24; DB 19; Length 1304;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1358 TTTTGTGAAAAA 1381
DB 1266 TTTTGTGAAAAA 1289
RESULT 25
AAF84329
ID AAF84329 standard; cDNA; 1311 BP.
XX
XX AAF84329;
XX
XX 20-JUN-2001 (first entry)
XX
XX Barley iron absorption improver coding sequence #2.
XX
XX Barley; iron absorption improver; Ids6; Gramineae; ss.
XX
XX Hordeum vulgare.
XX
XX Key Location/Qualifiers
XX CDS 58..1041
XX FT /*tag= a
XX FT /product= "Iron absorption improver"
XX
XX JP2001017181-A.
XX
XX 23-JAN-2001.
XX
XX 09-JUL-1999; 99JP-0196155.
XX
XX 09-JUL-1999; 99JP-0196155.
XX
XX Query Match 1.7%; Score 24; DB 19; Length 1304;
XX Best Local Similarity 100.0%; Pred. No. 4.8;
XX Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ID AAV30555 standard; DNA; 1304 BP.
XX
AC AAV30555;
XX
XX 14-OCT-1998 (first entry)
XX
DE Strawberry chalcone synthesis enzyme gene sequence.
XX
KW Chalcone synthesis; enzyme; strawberry; Fragaria ananassa; anthocyanin;
KW transgenic plant; colour; ds.
XX
XX Fragaria x ananassa.
OS
XX JP09308487-A.
XX
XX 02-DEC-1997.
XX
XX 22-MAY-1996; 96JP-0127245.
XX
XX 22-MAY-1996; 96JP-0127245.
XX
XX (ISHI ) ISHIKAWAJIMA HARIMA HEAVY IND.
XX
XX WPI; 1998-070648/07.
XX
XX Chalcone synthetic enzyme gene - used to produce plants of high
XX anthocyanin productivity
XX
XX Claim 1; Page 2; 5pp; Japanese.
XX
XX This sequence represents the gene encoding a chalcone synthetic enzyme
XX from strawberries (Fragaria x ananassa). The gene is used to produce
XX plant cells with increased anthocyanin productivity. This can lead to
XX the production of transformed plants which differ in colour from
XX conventional plants of that species.
XX
XX Sequence 1304 BP; 330 A; 309 C; 330 G; 335 T; 0 other;
XX
Query Match 1.7%; Score 24; DB 19; Length 1304;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1358 TTTTGTGAAAAA 1381
DB 1266 TTTTGTGAAAAA 1289
RESULT 25
AAF84329
ID AAF84329 standard; cDNA; 1311 BP.
XX
XX AAF84329;
XX
XX 20-JUN-2001 (first entry)
XX
XX Barley iron absorption improver coding sequence #2.
XX
XX Barley; iron absorption improver; Ids6; Gramineae; ss.
XX
XX Hordeum vulgare.
XX
XX Key Location/Qualifiers
XX CDS 58..1041
XX FT /*tag= a
XX FT /product= "Iron absorption improver"
XX
XX JP2001017181-A.
XX
XX 23-JAN-2001.
XX
XX 09-JUL-1999; 99JP-0196155.
XX
XX 09-JUL-1999; 99JP-0196155.
XX
XX Query Match 1.7%; Score 24; DB 22; Length 1312;
XX Best Local Similarity 100.0%; Pred. No. 4.8;
XX Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ID AAF84327 standard; cDNA; 1312 BP.
XX
XX AAF84327;
XX
XX 20-JUN-2001 (first entry)
XX
XX Barley iron absorption improver coding sequence #1.
XX
XX Barley; iron absorption improver; Ids6; Gramineae; ss.
XX
XX Hordeum vulgare.
XX
XX Key Location/Qualifiers
XX CDS 59..1042
XX FT /*tag= a
XX FT /product= "Iron absorption improver"
XX
XX JP2001017181-A.
XX
XX 23-JAN-2001.
XX
XX 09-JUL-1999; 99JP-0196155.
XX
XX 09-JUL-1999; 99JP-0196155.
XX
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
XX WPI; 2001-285407/30.
XX
XX P-PSDB; AAB80999.
XX
XX New gene encoding a 36 kDa protein is useful as an absorption improver
XX for iron-deficient barley roots -
XX
XX Claim 4; Page 16; 28pp; Japanese.
XX
XX The present sequence is a coding sequence for an iron absorption
XX improver, (Ids6) for Gramineae. This sequence was derived from
XX iron-deficient barley roots.
XX
XX Sequence 1312 BP; 239 A; 398 C; 422 G; 253 T; 0 other;
XX
Query Match 1.7%; Score 24; DB 22; Length 1312;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1358 TTTTGTGAAAAA 1381
DB 1285 TTTTGTGAAAAA 1308
RESULT 26
AAF84327
ID AAF84327 standard; cDNA; 1312 BP.
XX
XX AAF84327;
XX
XX 20-JUN-2001 (first entry)
XX
XX Barley iron absorption improver coding sequence #1.
XX
XX Barley; iron absorption improver; Ids6; Gramineae; ss.
XX
XX Hordeum vulgare.
XX
XX Key Location/Qualifiers
XX CDS 59..1042
XX FT /*tag= a
XX FT /product= "Iron absorption improver"
XX
XX JP2001017181-A.
XX
XX 23-JAN-2001.
XX
XX 09-JUL-1999; 99JP-0196155.
XX
XX 09-JUL-1999; 99JP-0196155.
XX
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
XX WPI; 2001-285407/30.
XX
XX P-PSDB; AAB80999.
XX
XX New gene encoding a 36 kDa protein is useful as an absorption improver
XX for iron-deficient barley roots -
XX
XX Claim 4; Page 16; 28pp; Japanese.
XX
XX The present sequence is a coding sequence for an iron absorption
XX improver, (Ids6) for Gramineae. This sequence was derived from
XX iron-deficient barley roots.
XX
XX Sequence 1312 BP; 239 A; 398 C; 422 G; 253 T; 0 other;
XX
Query Match 1.7%; Score 24; DB 22; Length 1312;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1358 TTTTGTGAAAAA 1381
DB 1285 TTTTGTGAAAAA 1308

```

QY 1358 TTTTGTGAAAAA 1381
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1286 TTTTGTGAAAAA 1309

RESULT 27
AAC79701
ID AAC79701 standard; cDNA; 1560 BP.
XX AAC79701;
AC
XX
DT 12-FEB-2001 (first entry)
XX
DE Human secreted protein gene 21 SEQ ID NO:31.
XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antitumor; vulnerary; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
KW food additive; preservative; chromosome identification; cancer;
KW immune disorder; cardiovascular disorder; neurological disease;
KW wound healing; infectious disease; ss.
XX Homo sapiens.
OS
XX
XX
PN WO200058339-A2.
XX
XX 05-OCT-2000.
PD
XX
XX 22-MAR-2000; 2000WO-US07440.
PF
XX
XX 26-MAR-1999; 99US-0126503.
PR
XX 17-DEC-1999; 99US-0172409.
PR
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
PI
XX WPI: 2000-594637/56.
DR
XX P-PSDB; AAB44616.
DR
XX
XX Fifty nucleic acid molecules encoding human secreted proteins, useful
PT in the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases -
PT
XX Claim 1; Page 349-350; 410pp; English.
PS
XX
XX The polynucleotide sequences given in AAC79681 to AAC79730 encode the
CC human secreted proteins given in AAB44596 to AAB44645. AAB44646 to
CC AAB44693 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Examples of activities include:
CC cytostatic; immunosuppressive; neotropic; neuroprotective; antiviral;
CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
CC vulnerary; anticonvulsant; antidiabetic; antifungal; antiparasitic; and
CC cardiant. The polynucleotides and polypeptides are useful for preventing,
CC treating or ameliorating a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. The polypeptides
CC can also be used as a food additive or preservative to increase or
CC decrease storage capabilities. The polynucleotides are useful for
CC chromosome identification. They are also useful as probes for diagnosing
CC a disorder related to the female reproductive system, particularly breast
CC and/or ovarian cancer. They are also useful in the gene therapy of breast
CC and ovarian cancer. The nucleic acids, protein, antibodies, agonists and
CC antagonists from the present invention are useful in the diagnosis,
CC treatment and prevention of: cancer; immune disorders; cardiovascular
CC disorders; wound healing; neurological diseases; and infectious
CC diseases. AAC79672 to AAC79680 and AAB44595 represent sequences used in
CC the exemplification of the present invention.
XX
XX Sequence 1560 BP; 313 A; 400 C; 438 G; 407 T; 2 other;

Query Match 1.7%; Score 24; DB 21; Length 1560;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAA 1381
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1536 TTTTGTGAAAAA 1559

RESULT 28
AAC64785
ID AAC64785 standard; cDNA; 1648 BP.
XX AAC64785;
AC
XX
DT 28-FEB-2001 (first entry)
XX
DE Lemon acyl transferase encoding cDNA SEQ ID NO:21A.
XX
KW Strawberry; fruit flavour; biosynthetic pathway; aliphatic; thiolase;
KW aromatic ester; alcohol acyl transferase; alcohol dehydrogenase;
KW pyruvate decarboxylase; aminotransferase; esterase; alcohol; aldehyde;
KW alpha-keto acid; amino acid; fatty acid; acyl-CoA; processed food;
KW food additive; flavour; syrup; ice-cream; frozen dessert; yoghurt;
KW confectionery; flavouring; oral medication; vitamin; aroma; beverage;
KW alcohol; scent; fragrance; perfume; cosmetic; cleaning product;
KW aluminum salt; anti-perspirant; pharmaceutical; insect repellent; miticide;
KW insect pheromone; dye carrier; solvent; insect repellent; miticide;
KW scabicide; plasticiser; deodorant; ss.
XX
XX Citrus limon.
OS
XX
XX WO200032789-A1.
PN
XX
XX 08-JUN-2000.
PD
XX
XX 02-DEC-1999; 99WO-NL00737.
PF
XX
XX 02-DEC-1998; 98EP-0204018.
PR
XX 12-MAR-1999; 99EP-0200739.
PR
XX
XX (CPRO-) CPRO-DLO CENT PLANTENVEREDELINGS REPROD.
PA
XX
PI Aharoni A, Luecker J, Verhoeven HA, Van Tunen AJ, O'Connell AP;
PI
XX WPI: 2000-412335/35.
DR
XX P-PSDB; AAB36459.
DR
XX
XX A new DNA sequence encoding a polypeptide with alcohol acyl transferase
PT activity for producing and regulating aromatic and/or aliphatic ester
PT formation in microorganisms, plant cells or plants -
PT
XX Example 5; Page 116-117; 163pp; English.
PS
XX
XX The present invention describes nucleotide sequences with thiolase,
CC alcohol acyl transferase, alcohol dehydrogenase, pyruvate decarboxylase,
CC aminotransferase and esterase activities, which are involved in the
CC biosynthetic pathway for aliphatic and/or aromatic ester production in
CC fruit. The nucleotide sequences can be inserted into the genome of a
CC fruit-producing plant to regulate aliphatic and/or aromatic ester
CC formation. Aromatic and/or aliphatic esters in microorganisms, plant
CC cells or plants are produced by inserting thiolase, alcohol acyl
CC transferase, alcohol dehydrogenase, pyruvate decarboxylase,
CC aminotransferase and esterase nucleotide sequences into the genome and
CC feeding the microorganism or plant with alcohol, aldehydes, alpha-keto
CC acids, or amino acids and fatty acids, and acyl-CoA. The nucleotides and
CC their proteins can be used in the processed food industry as food
CC additives to enhance the flavour of syrups, ice-creams, frozen desserts,
CC yoghurts and confectionery. They are used: as flavouring agents for oral
CC medications and vitamins; provide flavour and aroma in beverages,
CC including alcohol; enhance or reduce fruit flavour, aroma, fragrance or
CC scent; enhance the flavour or aroma of natural, synthetic or artificial
CC products; for the production of novel combinations of artificial flavour

CC substances; as antibacterial or anti-fungal agents; as fragrance or
 CC perfumes in cosmetics, creams, sun-protectant products, hair
 CC conditioners, lengthening agents and fixatives in perfumes, suspension
 CC aids for aluminum salts in anti-perspirant pharmaceuticals, cleaning
 CC products, personal care products and animal care products; as
 CC disinfectant additives; as degreasing solvents for electronics; as
 CC insect pheromones; and as dye carriers, solvents, insect repellents,
 CC miticides, scabicides, plasticisers and deodorants. The present sequence
 CC encodes lemon acyl transferase, from the present invention.
 XX
 XX Sequence 1648 BP; 464 A; 344 C; 354 G; 486 T; 0 other;
 SQ

Query Match 1.7%; Score 24; DB 21; Length 1648;
 Best Local Similarity 100.0%; Pred. No. 4.7;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAA 1381
 |||||
 DB 1624 TTTTGTGAAAAA 1647
 |||||

RESULT 29
 AAC81068
 ID AAC81068 standard; cDNA: 1727 BP.
 XX
 AC AAC81068;
 XX
 DT 13-FEB-2001 (first entry)
 XX
 DE Human secreted protein cDNA sequence #41.
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200063230-A2.
 XX
 PD 26-OCT-2000.
 XX
 PF 23-MAR-2000; 2000WO-US07677.
 XX
 PR 26-MAR-1999; 99US-0126601.
 PR 17-SEP-1999; 99US-0154373.
 PR 14-JAN-2000; 2000US-0176064.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Komatsoulis G;
 XX
 DR WPI: 2000-647515/62.
 DR P-PSDB; AAB45266.
 XX
 PT New nucleic acid molecules encoding 49 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives -
 XX
 XX Claim 1; Page 363-364; 402pp; English.
 XX
 CC The present invention relates to the isolation of genes encoding
 CC 49 human secreted proteins. The genes can be used to
 CC generate fusion proteins by linking to the gene for the human
 CC immunoglobulin G Fc portion for increasing the stability of
 CC the fusion protein as compared to the human protein only. The genes and
 CC proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated
 CC from a range of human tissues disclosed in the specification. The
 CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer, and other cancers of the adrenal gland, bone, bone

CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections. The present sequence is a cDNA encoding a protein
 CC of the invention.
 XX
 XX Sequence 1727 BP; 441 A; 389 C; 402 G; 494 T; 1 other;
 SQ

Query Match 1.7%; Score 24; DB 21; Length 1727;
 Best Local Similarity 100.0%; Pred. No. 4.6;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAA 1381
 |||||
 DB 1690 TTTTGTGAAAAA 1713
 |||||

RESULT 30
 AAD15431
 ID AAD15431 standard; cDNA: 1807 BP.
 XX
 AC AAD15431;
 XX
 DT 15-NOV-2001 (first entry)
 XX
 DE Wheat protoporphyrinogen oxidase-1 (protox-1) partial cDNA.
 KW Protoporphyrinogen oxidase-1; Protox-1; transgenic plant; crop; soybean;
 KW sugarcane; barley; cotton; tobacco; plant cell transformation; pesticide;
 KW tissue culture; herbicide-tolerant; protoporphyrin IX; wheat; ss.
 XX
 OS Triticum aestivum.
 XX
 FH Key Location/Qualifiers
 FT CDS 3..1589
 FT /*tag= a
 FT /product= "Wheat Protox-1 protein"
 FT /note= "CDS does not include start codon"
 FT /partial
 XX
 PN US2001016956-A1.
 XX
 PD 23-AUG-2001.
 XX
 PF 05-DEC-2000; 2000US-0730525.
 XX
 PR 28-FEB-1996; 96US-0012705.
 PR 28-FEB-1996; 96US-0013612.
 PR 21-JUN-1996; 96US-0020003.
 PR 11-MAR-1998; 98US-0126430.
 PR 16-JUN-1994; 94US-0261198.
 PR 06-JUN-1995; 95US-0472028.
 PR 28-FEB-1997; 97US-0808931.
 PR 30-MAR-1998; 98US-0050603.
 PR 13-APR-1998; 98US-0059164.
 PR 22-JUN-1998; 98US-0102419.
 XX
 PA (WARD/) WARD E R.
 PA (VOLK/) VOLK RATH S L.
 PA (JOHN/) JOHNSON M A.
 PA (PORT/) POTTER S L.
 XX
 PI Ward ER, Vollrath SL, Johnson MA, Potter SL;
 DR WPI: 2001-529188/58.
 DR P-PSDB; AAE08752.
 XX
 PT Novel shuffled DNA molecule obtained by shuffling template DNA molecule
 PT having protox enzyme activity, encodes protox enzyme having enhanced

PT tolerance to herbicide that inhibits protox activity encoded by
PT template DNA

PS Example 1; Page 66-68; 98pp; English.

XX The invention relates to protoporphyrinogen oxidase (Protox) enzymes and
XX their corresponding cDNA molecules. Protox enzyme catalyses the oxidation
CC of protoporphyrinogen IX to protoporphyrin IX. The invention also relates
CC to herbicide-tolerant plants as well as methods for tissue culture
CC selection and herbicide application based on the herbicide-tolerant forms
CC of protox. The invention is useful for producing a protox enzyme having
CC enhanced tolerance to herbicide that inhibits the protox activity encoded
CC by a template DNA molecule. A shuffled (mutagenised) DNA molecule
CC encoding the herbicide resistant protox enzyme is genetically engineered
CC for optimal expression in a crop plant, and thus is useful for producing
CC plants e.g. sugarcane, soybean, barley, cotton, tobacco, etc. or its
CC progeny, plant tissues, plant seeds tolerant to herbicides that inhibit
CC the naturally occurring protox activity in these plants. Transgenic seeds
CC and plants produced using shuffled DNA molecules are used for breeding
CC improved plant lines that e.g. increase the effectiveness of conventional
CC methods such as herbicide or pesticide treatment or allow to dispense
CC with the methods due to their modified genetic properties. A shuffled DNA
CC molecule encoding altered protox tolerant to a protox inhibitor is also
CC used as selectable markers in plant cell transformation methods.
CC The present cDNA sequence encodes wheat protoporphyrinogen oxidase-1
CC (protox-1) partial protein.

XX Sequence 1807 BP; 421 A; 463 C; 516 G; 407 T; 0 other;

Query Match 1.7%; Score 24; DB 22; Length 1807;

Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAAATAAAAAAAAAA 1381

DB 1782 TTTTGTGAAAAAATAAAAAAAAAA 1805

RESULT 31

AAV04308

ID AAV04308 standard; cDNA; 1811 BP.

XX AAV04308;

DT 20-APR-1998 (first entry)

XX Wheat protox-1 cDNA.

XX protoporphyrinogen oxidase-1; protox-1; promoter; wheat;
KW herbicide resistance; breeding programme; probe; gene isolation;
KW genomic mapping; ss.

XX Triticum aestivum.

XX Key Location/Qualifiers

FT CDS 3..1589

FT /*tag= a

FT /product= protox-1

XX WO9732028-A1.

XX 04-SEP-1997.

XX 27-FEB-1997; 97WO-US03343.

XX 21-JUN-1996; 96US-0020003.

XX 28-FEB-1996; 96US-0012705.

XX 28-FEB-1996; 96US-0013612.

XX (NOVS) NOVARTIS AG.

XX Johnson MA, Volrath SL, Ward ER;

XX

DR WPI; 1997-489209/45.
DR P-PSDB; AAW41607.

XX DNA containing a plant proto-porphyrinogen oxidase gene promoter -
PT optionally linked to a heterologous gene, especially to express
PT herbicide-resistant enzymes, and plants containing such constructs
XX Disclosure; Pages 61-64; 114pp; English.

XX The present sequence encodes wheat protoporphyrinogen oxidase-1
CC (protox-1).

CC The protox-1 promoter can be used to express herbicide resistant
CC enzymes, specifically protox, i.e. a plant tissue, plant or progeny
CC containing a chimeric gene of the promoter and a heterologous
CC coding sequence. The plant can also be used in breeding programmes.
CC Also hybridising fragments of the protox coding sequence can be
CC used as probes, e.g. to isolate related genes or for genomic
CC mapping.

XX Sequence 1811 BP; 423 A; 463 C; 517 G; 408 T; 0 other;

Query Match 1.7%; Score 24; DB 18; Length 1811;

Best Local Similarity 100.0%; Pred. No. 4.6;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAAATAAAAAAAAAA 1381

DB 1786 TTTTGTGAAAAAATAAAAAAAAAA 1809

RESULT 32

AAT86121

ID AAT86121 standard; cDNA; 1811 BP.

XX AAT86121;

XX 01-MAR-1998 (first entry)

XX Wheat protoporphyrinogen oxidase (protox-1) cDNA clone pWDC-13.

XX Protox-1a; protoporphyrinogen oxidase; inhibitor; wheat;

KW herbicide tolerance; herbicide resistance; transgenic plant; ss.

XX Triticum aestivum cv. Kanzler.

XX Key Location/Qualifiers

FT CDS 3..1589

FT /*tag= a

XX WO9732011-A1.

XX 04-SEP-1997.

XX 27-FEB-1997; 97WO-US03313.

XX 21-JUN-1996; 96US-0020003.

XX 28-FEB-1996; 96US-0012705.

XX 28-FEB-1996; 96US-0013612.

XX (NOVS) NOVARTIS AG.

XX Heifetz PB, Johnson MA, Potter SL, Volrath SL, Ward ER;

XX WPI; 1997-448683/41.

DR P-PSDB; AAW25738.

XX New DNA encoding plant protoporphyrinogen oxidase enzyme - and
PT herbicide resistant mutants; useful to prepare plants resistant to
PT herbicide which therefore kills undesired vegetation only

XX Claim 4; Page 120-123; 196pp; English.

XX This cDNA clone codes for wheat protoporphyrinogen oxidase

CC

CC (Protox-1a, see AAW25738). It was isolated from a cDNA library using
 CC a previously isolated maize Protox-1 cDNA clone (see AAT86128) as
 CC probe. Wheat Protox-1a in pBluescript SK vector has been deposited
 CC as PWDG-13 (NRRL AAB21545). The isolated cDNA can be mutated so that
 CC it encodes a modified protox enzyme that is resistant to protox
 CC inhibitors and hence is herbicide tolerant. Plants, especially
 CC crop plants, may be engineered for resistance to protox inhibitors
 CC via mutation of the native protox gene to a resistant form, or they
 CC may be transformed with a gene encoding an inhibitor-resistant form
 CC of a plant protox enzyme, including claimed forms from wheat,
 CC soybean, cotton, sugarbeet, oilseed rape, rice and sorghum (see
 CC AAW25738-45). Application of herbicide therefore kills undesired
 CC vegetation only. Plant protox nucleic acids can also be used as
 CC probes and PCR primers, as selectable markers in plant cell
 CC transformation methods, and for recombinant production of protox
 CC enzymes in host cells.

XX Sequence 1811 BP; 423 A; 463 C; 517 G; 408 T; 0 other;

Query Match 1.7%; Score 24; DB 18; Length 1811;
 Best Local Similarity 100.0%; Pred. No. 4.6;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAAATAAAAAA 1381
 |||||
 DB 1786 TTTTGTGAAAAAATAAAAAA 1809

RESULT 33

AAAF6575

ID AAF6575 standard; cDNA; 1811 BP.

XX AC AAF6575;

DT 16-MAY-2001 (first entry)

XX Wheat protoporphyrinogen oxidase coding sequence SEQ ID NO: 9.

XX Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
 KW soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein; ss.

XX Triticum aestivum.

XX WO200112825-A1.

XX 22-FEB-2001.

XX 30-JUN-2000; 2000WO-EP06127.

XX 13-AUG-1999; 99US-0373691.

XX (SYNG-) SYNGENTA PARTICIPATIONS AG.

XX Johnson MA, Volrath SL, Heifetz PB, Law MD;

XX WPI; 2001-234914/24.

XX P-PSDB; AAB72905.

XX Plant DNA molecules encoding herbicide-tolerant forms of
 PT protoporphyrinogen oxidase which are useful for rationally designing
 PT new inhibitory herbicides and for producing herbicide-tolerant
 PT transgenic plants and seeds -

XX Claim 2; Page 161-165; 228pp; English.

XX The present invention provides the protein and coding sequences of a
 CC number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
 CC sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
 CC oxidase (protox) enzyme. Examples of these mutants are shown in
 CC AAB72920-AAB72926. They are useful as they enable the production of
 CC herbicide-tolerant plants and seeds. The present sequence is a protox
 CC coding sequence.

XX

SQ Sequence 1811 BP; 423 A; 463 C; 517 G; 408 T; 0 other;

Query Match 1.7%; Score 24; DB 22; Length 1811;
 Best Local Similarity 100.0%; Pred. No. 4.6;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAAATAAAAAA 1381

|||||

DB 1786 TTTTGTGAAAAAATAAAAAA 1809

RESULT 34

RAD17335

ID AAD17335 standard; cDNA; 1811 BP.

XX AC AAD17335;

DT 29-NOV-2001 (first entry)

XX Wheat protoporphyrinogen oxidase (protox-1) cDNA.

XX Protoporphyrinogen oxidase; protox-1; herbicide; transgenic plant;
 KW wheat; tissue culture; ss.

XX Triticum aestivum.

XX Key Location/Qualifiers

XX CDS 3..1589

XX FT /*tag= a

XX FT /*product= "Wheat protox-1 protein"

XX FT /*note= "CDS does not include start codon"

XX FT /*partial

XX WO200168826-A2.

XX 20-SEP-2001.

XX 12-MAR-2001; 2001WO-EP02753.

XX 14-MAR-2000; 2000US-0525152.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Fonne-Pfister R, De Marco A, Volrath SL;

XX WPI; 2001-557932/62.

XX P-PSDB; AAE10224.

XX An isolated maize DNA molecule encoding a protoporphyrinogen oxidase
 PT (protox) which are useful for rationally designing new inhibitory
 PT herbicides and for producing herbicide-tolerant transgenic plants and
 PT seeds -

XX Example 1; Page 136-139; 183pp; English.

XX The invention relates to DNA molecules encoding protoporphyrinogen
 CC oxidase (protox) enzyme and herbicide-tolerant forms of the enzyme. The
 CC invention further relates to herbicide-tolerant plants as well as methods
 CC for tissue culture selection and herbicide application based on these
 CC herbicide-tolerant forms of protox. The nucleic acids can be used to
 CC produce herbicide-tolerant transgenic plants and seeds. Recombinantly
 CC produced eukaryotic protox enzyme is useful in an in vitro assay to
 CC screen known herbicidal chemicals whose target has not been identified
 CC to determine if they inhibit protox. Recombinantly produced eukaryotic
 CC protox enzyme may also be used in an assay to identify inhibitor-
 CC resistant protox mutants. Alternatively, recombinantly produced protox
 CC enzyme may be used to further characterise its association with known
 CC inhibitors in order to rationally design new inhibitory herbicides as
 CC well as herbicide tolerant forms of the enzyme. Genes encoding altered
 CC protox enzymes can be used as selectable markers in plant transformation
 CC methods. The present sequence is a cDNA encoding wheat protox-1 protein.

XX Sequence 1811 BP; 423 A; 463 C; 517 G; 408 T; 0 other;


```
Query Match      1.7%; Score 24; DB 21; Length 2010;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAA... 1381
Db 1985 TTTTGTGAAAAA... 2008

RESULT 37
AAZ52564
ID AAZ52564 standard; cDNA; 2080 BP.
XX
AC AAZ52564;
XX
DT 29-FEB-2000 (first entry)
DE
DE Human secreted protein clone yk38_1 nucleotide sequence SEQ ID NO:179.
XX
KW Human; secreted protein; immunostimulatory; haemostatic; cytokine;
KW proliferative; differentiative; chemotactic; chemokinetic; vaccine;
KW thrombolytic; antiinflammatory; cytostatic; immunosuppressive;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO9958642-A2.
XX
PD 18-NOV-1999.
XX
PF 14-MAY-1999; 99WO-US10843.
XX
PR 14-MAY-1998; 98US-0085472.
PR 17-AUG-1998; 98US-0096824.
PR 11-SEP-1998; 98US-0099843.
PR 11-SEP-1998; 98US-0099950.
PR 15-SEP-1998; 98US-0100424.
PR 29-SEP-1998; 98US-0102329.
PR 09-OCT-1998; 98US-0103615.
PR 11-DEC-1998; 98US-0111799.
PR 14-DEC-1998; 98US-0112159.
PR 31-DEC-1998; 98US-0114415.
PR 10-FEB-1999; 99US-0248059.
PR 06-APR-1999; 99US-0287150.
PR 13-MAY-1999; 99US-0311021.
XX
PA (GEM) GENETICS INST INC.
XX
XX Wong GG, Clark HF, Fecthel K, Agostino MJ;
XX WPI; 2000-053095/04.
XX P-PSDB; AAY73479.
XX
XX Novel polynucleotides and proteins having biological activities which
XX make them suitable for treating, preventing or ameliorating medical
XX conditions in humans or animals -
XX
XX Claim 188; Page 692-693; 730pp; English.
XX
XX The present invention describes human secreted proteins encoded by
XX polynucleotides obtained from adult testes, foetal brain, adult brain,
XX brain (foetal and adult), foetal kidney, adult spleen, and adult thymus
XX cDNA libraries. The polynucleotides and proteins are predicted to have
XX biological activities which would make them suitable for treating,
XX preventing or ameliorating medical conditions in humans and animals.
XX Suggested activities include nutritional activity, cytokine and cell
XX proliferation/differentiation activity, immune stimulating (e.g. as
XX vaccines) or suppressing activity, haematopoiesis regulating activity,
XX tissue growth activity, activin/inhibin activity, chemotactic/
XX chemokinetic activity, haemostatic and thrombolytic activity, receptor/
XX ligand activity, anti-inflammatory activity, cadherin/tumour invasion
XX suppressor activity, and tumour inhibition activity. The polynucleotides
XX
```

```
CC are also stated to be useful for gene therapy. Therapeutic compositions
CC are also presently valuable for veterinary applications. AA52475 to
CC AA52581 encode human secreted proteins, and AAY73390 to AAY73500
CC represent human secreted proteins, given in the present invention.
XX
XX Sequence 2080 BP; 494 A; 502 C; 526 G; 558 T; 0 other;

Query Match      1.7%; Score 24; DB 21; Length 2080;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAA... 1381
Db 2044 TTTTGTGAAAAA... 2067

RESULT 38
AAV59751
ID AAV59751 standard; DNA; 2082 BP.
XX
AC AAV59751;
XX
DT 19-JAN-1999 (first entry)
DE
DE Human secreted protein gene 95 clone HMDAA61.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; testenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX Homo sapiens.
XX
XX WO9839448-A2.
XX
XX 11-SEP-1998.
XX
XX 06-MAR-1998; 98WO-US04493.
XX
XX 02-OCT-1997; 97US-0061060.
XX 07-MAR-1997; 97US-0038621.
XX 07-MAR-1997; 97US-0040161.
XX 07-MAR-1997; 97US-0040162.
XX 07-MAR-1997; 97US-0040163.
XX 07-MAR-1997; 97US-0040333.
XX 07-MAR-1997; 97US-0040334.
XX 07-MAR-1997; 97US-0040336.
XX 07-MAR-1997; 97US-0040626.
XX 11-APR-1997; 97US-0043311.
XX 11-APR-1997; 97US-0043312.
XX 11-APR-1997; 97US-0043313.
XX 11-APR-1997; 97US-0043314.
XX 11-APR-1997; 97US-0043568.
XX 11-APR-1997; 97US-0043569.
XX 11-APR-1997; 97US-0043576.
XX 11-APR-1997; 97US-0043578.
XX 11-APR-1997; 97US-0043580.
XX 11-APR-1997; 97US-0043669.
XX 11-APR-1997; 97US-0043670.
XX 11-APR-1997; 97US-0043671.
XX 11-APR-1997; 97US-0043672.
XX 11-APR-1997; 97US-0043674.
XX 23-MAY-1997; 97US-0047492.
XX 23-MAY-1997; 97US-0047500.
XX 23-MAY-1997; 97US-0047501.
XX 23-MAY-1997; 97US-0047502.
XX 23-MAY-1997; 97US-0047503.
XX 23-MAY-1997; 97US-0047501.
XX 23-MAY-1997; 97US-0047582.
```


XX WO2000077022-A1.
 XX 21-DEC-2000.
 XX 01-JUN-2000; 2000WO-US15136.
 XX 11-JUN-1999; 99US-0138629.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM, Komatsoulis GA;
 XX WPI; 2001-367020/38.
 DR P-PSDB; AAE03898.
 XX
 XX Nucleic acids encoding 50 human secreted polypeptides, useful for
 PT preventing, diagnosing and/or treating diseases, e.g. Parkinson's
 PT disease, botulism, cancers and Scimitar syndrome -
 XX
 PS Claim 1: Page 485-486; 614pp; English.
 XX
 CC AAD08345-AAD08394 represent cDNAs corresponding to 50 human secreted
 CC protein genes and AAE03898-AAE03947 represent the proteins they encode.
 CC AAE03948-AAE03996 represent human secreted protein fragments or variants.
 CC The genes and their secreted proteins are useful for preventing,
 CC treating or ameliorating medical conditions, e.g., by protein or gene
 CC therapy. Pathological conditions can be diagnosed by determining the
 CC amount of the new protein in a sample or by determining the presence of
 CC mutations in the new genes. Specific uses are described for each of the
 CC 50 genes, based on the tissues in which they are most highly expressed,
 CC and include developing products for the diagnosis or treatment of
 CC proliferative disorders, cancer, tumours, foetal and developmental
 CC abnormalities, haematopoietic disorders, diseases of the immune system,
 CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
 CC allergies, neurological disorders (e.g., Alzheimer's disease,
 CC Parkinson's disease), cognitive disorders, schizophrenia, asthma,
 CC skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
 CC cardiovascular disorders, angiogenic disorders, kidney disorders,
 CC gastrointestinal disorders, pregnancy-related disorders, endocrine
 CC disorders, and infections. The proteins can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin aging due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues, to identify their
 CC cognate ligands or binding partners, and in chemotaxis, and can be used
 CC as a food additive or preservative to modify storage properties.
 CC Antibodies specific for a protein of the invention can be used in
 CC alleviating symptoms associated with the disorders mentioned above, and
 CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
 CC immunosorbent assay (ELISA). The present sequence represents a human
 CC secreted protein-encoding cDNA of the invention.
 XX
 SQ Sequence 3350 BP; 908 A; 756 C; 695 G; 975 T; 16 other;
 Query Match 1.7%; Score 24; DB 22; Length 3350;
 Best Local Similarity 100.0%; Pred. No. 4.2;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1358 TTTTGTGAAAAAATAAAAAA 1381
 DB 3283 TTTTGTGAAAAAATAAAAAA 3306
 RESULT 40
 ABA19864
 ID ABA19864 standard; DNA; 5828 BP.
 XX
 AC ABA19864;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 12195.
 XX

KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
 KW antiparkinsonian; antiskickling; antianaemic; antiarthritic; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 XX Homo sapiens.
 OS
 XX WO200159063-A2.
 PN
 XX 16-AUG-2001.
 PD
 XX 17-JAN-2001; 2001WO-US01334.
 PF
 XX
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.

```
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234224.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249254.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.

PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2000US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system
XX cancers and metastases
XX Disclosure; SEQ ID NO 12195; 1701pp + Sequence Listing; English.
XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins
XX (ABBI4678-ABBI8001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune;
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemias;
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX SQ Sequence 5828 BP; 1698 A; 1052 C; 1193 G; 1885 T; 0 other;

Query Match 1.7%; Score 24; DB 22; Length 5828;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAA 1381
Db 1153 TTTTGTGAAAAA 1176

RESULT 41
ABAI9329/c
XX ID ABAI9329 standard; DNA; 6843 BP.
XX AC ABAI9329;
XX XX
XX DT 23-JAN-2002 (first entry)
XX DE Human nervous system related polynucleotide SEQ ID NO 11660.
XX KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
XX antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
XX antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
XX antiparasitic; cardiant; immune disorder; cardiovascular disorder;
XX neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX OS Homo sapiens.
XX XX
XX PN WO200159063-A2.
XX PD 16-AUG-2001.
XX XX
```

```
PR 17-JAN-2001; 2001WO-US01334.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227709.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244674.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT
```


PR 20-OCT-2000; 2000US-0241826.
 PR 20-OCT-2000; 2000US-0242221.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246612.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250391.
 PR 01-DEC-2000; 2000US-0251160.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-541565/60.
 XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system
 PT cancers and metastases -
 XX
 PS Disclosure; SEQ ID NO 12196; 1701pp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (AB11004-AB21534) and proteins
 CC (AB114678-AB118001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune

CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 9928 BP; 2916 A; 1786 C; 1965 G; 3261 T; 0 other;
 Query Match 1.7%; Score 24; DB 22; Length 9928;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1358 TTTTGTGAAAAA 1381
 Db 5283 TTTTGTGAAAAA 5306
 RESULT 43
 ABL32894
 ID ABL32894 standard; DNA; 11729 BP.
 XX AC ABL32894;
 XX DT 26-MAR-2002 (first entry)
 XX DE Human immune system associated gene SEQ ID NO: 867.
 KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianemic; cytostatic; nontropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX OS Homo sapiens.
 XX WO200200928-A2.
 XX PD 03-JAN-2002.
 XX PF 02-JUL-2001; 2001WO-EP07537.
 XX PR 30-JUN-2000; 2000DE-1032529.
 XX PR 01-SEP-2000; 2000DE-1043826.
 XX PA (EPIG-) EPIGENOMICS AG.
 XX PI Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2002-130909/17.
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX
 PS Claim 1; SEQ ID NO 867; 32pp + Sequence Listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX SQ Sequence 11729 BP; 3017 A; 201 C; 3016 G; 5495 T; 0 other;

Query Match	1.7%;	Score 24;	DB 24;	Length 11729;
Best Local Similarity	100.0%;	Pred. No. 3.5;		
Matches 24;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
<hr/>				
QY	1358	TTTTTTGAAAAA	AAAAAAAAAAAA	1381
Db	1904	TTTTTTGAAAAA	AAAAAAAAAAAA	1927
<hr/>				
RESULT 44				
ID ABAI7368				
XX	ABAI7368 standard; DNA; 32223 BP.			
XX	AC ABAI7368;			
XX				
XX	23-JAN-2002 (first entry)			
XX				
DE	Human nervous system related polynucleotide SEQ ID NO 9699.			
XX				
KW	Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;			
KW	immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;			
KW	antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;			
KW	antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;			
KW	antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;			
KW	antiparasitic; cardiant; immune disorder; cardiovascular disorder;			
KW	neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.			
XX				
OS	Homo sapiens.			
XX				
PN	WO200159063-A2.			
XX				
PD	16-AUG-2001.			
XX				
PF	17-JAN-2001; 2001WO-US01334.			
XX				
PR	31-JAN-2000; 2000US-0179065.			
PR	04-FEB-2000; 2000US-0180628.			
PR	24-FEB-2000; 2000US-0184664.			
PR	02-MAR-2000; 2000US-0186350.			
PR	16-MAR-2000; 2000US-0189874.			
PR	17-MAR-2000; 2000US-0190076.			
PR	18-APR-2000; 2000US-0198123.			
PR	19-MAY-2000; 2000US-0205515.			
PR	07-JUN-2000; 2000US-0209467.			
PR	28-JUN-2000; 2000US-0214886.			
PR	30-JUN-2000; 2000US-0215135.			
PR	07-JUL-2000; 2000US-0216647.			
PR	07-JUL-2000; 2000US-0216880.			
PR	11-JUL-2000; 2000US-0217487.			
PR	11-JUL-2000; 2000US-0217496.			
PR	14-JUL-2000; 2000US-0218290.			
PR	26-JUL-2000; 2000US-0220963.			
PR	26-JUL-2000; 2000US-0220964.			
PR	14-AUG-2000; 2000US-0224518.			
PR	14-AUG-2000; 2000US-0224519.			
PR	14-AUG-2000; 2000US-0225213.			
PR	14-AUG-2000; 2000US-0225214.			
PR	14-AUG-2000; 2000US-0225266.			
PR	14-AUG-2000; 2000US-0225267.			
PR	14-AUG-2000; 2000US-0225268.			
PR	14-AUG-2000; 2000US-0225270.			
PR	14-AUG-2000; 2000US-0225447.			
PR	14-AUG-2000; 2000US-0225757.			
PR	14-AUG-2000; 2000US-0225758.			
PR	14-AUG-2000; 2000US-0225759.			
PR	18-AUG-2000; 2000US-0226279.			
PR	22-AUG-2000; 2000US-0226681.			
PR	22-AUG-2000; 2000US-0226868.			
PR	22-AUG-2000; 2000US-0227182.			
PR	23-AUG-2000; 2000US-0227009.			
PR	30-AUG-2000; 2000US-0228924.			
PR	01-SEP-2000; 2000US-0229287.			
PR	01-SEP-2000; 2000US-0229343.			
PR	01-SEP-2000; 2000US-0229344.			
PR	01-SEP-2000; 2000US-0229345.			
PR	05-SEP-2000; 2000US-0229509.			
PR	05-SEP-2000; 2000US-0229513.			
PR	06-SEP-2000; 2000US-0230437.			
PR	06-SEP-2000; 2000US-0230438.			
PR	08-SEP-2000; 2000US-0231242.			
PR	08-SEP-2000; 2000US-0231243.			
PR	08-SEP-2000; 2000US-0231244.			
PR	08-SEP-2000; 2000US-0231413.			
PR	08-SEP-2000; 2000US-0231414.			
PR	08-SEP-2000; 2000US-0232080.			
PR	08-SEP-2000; 2000US-0232081.			
PR	12-SEP-2000; 2000US-0231968.			
PR	14-SEP-2000; 2000US-0232397.			
PR	14-SEP-2000; 2000US-0232398.			
PR	14-SEP-2000; 2000US-0232399.			
PR	14-SEP-2000; 2000US-0232400.			
PR	14-SEP-2000; 2000US-0232401.			
PR	14-SEP-2000; 2000US-0233063.			
PR	14-SEP-2000; 2000US-0233064.			
PR	14-SEP-2000; 2000US-0233065.			
PR	21-SEP-2000; 2000US-0234223.			
PR	21-SEP-2000; 2000US-0234274.			
PR	25-SEP-2000; 2000US-0234997.			
PR	25-SEP-2000; 2000US-0234998.			
PR	26-SEP-2000; 2000US-0235484.			
PR	27-SEP-2000; 2000US-0235834.			
PR	27-SEP-2000; 2000US-0235836.			
PR	29-SEP-2000; 2000US-0236327.			
PR	29-SEP-2000; 2000US-0236367.			
PR	29-SEP-2000; 2000US-0236368.			
PR	29-SEP-2000; 2000US-0236369.			
PR	29-SEP-2000; 2000US-0236370.			
PR	29-SEP-2000; 2000US-0236370.			
PR	02-OCT-2000; 2000US-0236802.			
PR	02-OCT-2000; 2000US-0237037.			
PR	02-OCT-2000; 2000US-0237038.			
PR	02-OCT-2000; 2000US-0237039.			
PR	02-OCT-2000; 2000US-0237040.			
PR	13-OCT-2000; 2000US-0239935.			
PR	13-OCT-2000; 2000US-0239937.			
PR	20-OCT-2000; 2000US-0240960.			
PR	20-OCT-2000; 2000US-0241785.			
PR	20-OCT-2000; 2000US-0241786.			
PR	20-OCT-2000; 2000US-0241787.			
PR	20-OCT-2000; 2000US-0241808.			
PR	20-OCT-2000; 2000US-0241809.			
PR	20-OCT-2000; 2000US-0241826.			
PR	20-OCT-2000; 2000US-0242221.			
PR	01-NOV-2000; 2000US-0244617.			
PR	08-NOV-2000; 2000US-0246474.			
PR	08-NOV-2000; 2000US-0246475.			
PR	08-NOV-2000; 2000US-0246476.			
PR	08-NOV-2000; 2000US-0246477.			
PR	08-NOV-2000; 2000US-0246478.			
PR	08-NOV-2000; 2000US-0246523.			
PR	08-NOV-2000; 2000US-0246524.			
PR	08-NOV-2000; 2000US-0246525.			
PR	08-NOV-2000; 2000US-0246526.			
PR	08-NOV-2000; 2000US-0246527.			
PR	08-NOV-2000; 2000US-0246528.			
PR	08-NOV-2000; 2000US-0246532.			
PR	08-NOV-2000; 2000US-0246609.			
PR	08-NOV-2000; 2000US-0246610.			
PR	08-NOV-2000; 2000US-0246611.			
PR	08-NOV-2000; 2000US-0246613.			
PR	17-NOV-2000; 2000US-0249207.			
PR	17-NOV-2000; 2000US-0249208.			
PR	17-NOV-2000; 2000US-0249209.			
PR	17-NOV-2000; 2000US-0249210.			
PR	17-NOV-2000; 2000US-0249211.			
PR	17-NOV-2000; 2000US-0249212.			

PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
XX Disclosure; SEQ ID NO 9699; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (ABR14678-ABR18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 32223 BP; 9124 A; 6178 C; 6281 G; 10640 T; 0 other;
Query Match 1.7%; Score 24; DB 22; Length 32223;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1358 TTTTGTGAAAAAATAAAAAA 1381
|||||
DB 27551 TTTTGTGAAAAAATAAAAAA 27574
RESULT 45
ABA19863
ID ABA19863 standard; DNA; 32223 BP.
XX

AC ABA19863;
XX
XX 23-JAN-2002 (first entry)
XX
XX Human nervous system related polynucleotide SEQ ID NO 12194.
XX
XX Human; neotropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01334.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.

```
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system
XX cancers and metastases.
XX
XX Disclosure; SEQ ID NO 12194; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AB11004-AB21534) and proteins
XX (AB114678-AB18001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemias;
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3223 BP; 9124 A; 6178 C; 6281 G; 10640 T; 0 other;
XX
XX Query Match 1.7%; Score 24; DB 22; Length 3223;
XX Best Local Similarity 100.0%; Pred. No. 3.1;
XX Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1358 TTTTGTGAAAAA 1381
Db 27551 TTTTGTGAAAAA 27574
Search completed: November 5, 2002, 13:50:17
Job time : 372.876 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
OM nucleic - nucleic search, using sw model
Run on: November 5, 2002, 11:01:37 ; Search time 33.87 Seconds
(without alignments)
10015.359 Million cell updates/sec
Title: US-09-805-311-5
Perfect score: 1381
Sequence: 1 cgaccacgctcgccac.....ttgaaaaaiaaaaaaaaaa 1381
Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0
Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2.6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2.6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2.6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2.6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2.6/ptodata/1/ina/ECTUS.COMB.seq.*
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1381	100.0	1381	4	US-09-426-557-5
2	1267	91.7	1463	4	US-09-426-557-1
3	1216	88.1	1541	4	US-09-426-557-3
4	850	61.5	1478	4	US-09-426-557-7
5	27	2.0	1096	3	US-09-136-073-1
6	24	1.7	1811	2	US-08-808-931-9
7	24	1.7	1811	3	US-08-808-323-9
8	24	1.7	1811	3	US-09-050-603A-9
9	24	1.7	1811	3	US-09-102-420B-9
10	24	1.7	1811	4	US-09-497-698-9
11	23	1.7	3138	3	US-09-234-332-5
12	22	1.6	1051	2	US-08-865-273-1
13	22	1.6	1051	4	US-09-385-174-1
14	22	1.6	1067	4	US-09-045-193-1
15	22	1.6	1140	2	US-08-698-805-5
16	22	1.6	1193	4	US-09-347-798-1
17	22	1.6	1493	2	US-08-820-170A-9
18	22	1.6	1493	3	US-09-055-699-9
19	22	1.6	1493	4	US-09-273-565-9
20	22	1.6	1493	4	US-09-565-538-9
21	22	1.6	1745	1	US-08-464-523B-1
22	22	1.6	3855	4	US-08-974-549A-4
23	22	1.6	4029	3	US-08-851-843A-173
24	22	1.6	4029	4	US-08-974-549A-292
25	22	1.6	4029	4	US-08-854-050-173
26	22	1.6	4029	4	US-09-430-323-173
27	21	1.5	40	4	US-09-306-230-41

c	28	21	1.5	360	4	US-09-227-357-121	Sequence 121, Appl
	29	21	1.5	518	1	US-08-485-284A-2	Sequence 2, Appl1
	30	21	1.5	700	1	US-07-846-992-1	Sequence 1, Appl1
	31	21	1.5	700	1	US-08-469-555-1	Sequence 1, Appl1
	32	21	1.5	720	3	US-08-946-026-14	Sequence 14, Appl
	33	21	1.5	855	6	5185441-40	Patent No. 5185441
	34	21	1.5	855	6	5223394-3	Patent No. 5223394
	35	21	1.5	863	1	US-07-940-861-11	Sequence 11, Appl
	36	21	1.5	863	1	US-08-459-512-11	Sequence 11, Appl
	37	21	1.5	863	2	US-08-459-657-11	Sequence 11, Appl
	38	21	1.5	863	2	US-08-460-132-11	Sequence 11, Appl
	39	21	1.5	863	5	PCT-US92-02050-11	Sequence 11, Appl
	40	21	1.5	863	6	5185441-35	Patent No. 5185441
	41	21	1.5	863	6	5223394-5	Patent No. 5223394
	42	21	1.5	880	1	US-08-616-368A-7	Sequence 7, Appl1
	43	21	1.5	880	3	US-09-054-298-7	Sequence 7, Appl1
	44	21	1.5	880	4	US-08-818-655-7	Sequence 7, Appl1
	45	21	1.5	1009	6	5223394-8	Patent No. 5223394

ALIGNMENTS

RESULT 1
US-09-426-557-5
; Sequence 5, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/PEN-1 Orthologues and Uses
; FILE REFERENCE: Thereof
; CURRENT APPLICATION NUMBER: US/09/426.557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112.332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37)...(1173)
US-09-426-557-5

Query Match	100.0%	Score 1381;	DB 4;	Length 1381;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1381;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGACCCACGCTCGGGCCACAGCCGCCGACAGATGGCGATGGCGATCAAGGGTTTGACGAAA	60	
Db	1	CGACCCACGCTCGGGCCACAGCCGCCGACAGATGGCGATGGCGATCAAGGGTTTGACGAAA	60	
QY	61	CTGCTGGCGGCAATGCGCCCAAGCGGATGAAGGAGCAGAGATTCGAGAGCTACTTCGGC	120	
Db	61	CTGCTGGCGGCAATGCGCCCAAGCGGATGAAGGAGCAGAGATTCGAGAGCTACTTCGGC	120	
QY	121	CGCAAAATCGCGCTCGACCCAGCATATACCAAGTTCCCTGATGTAGTTGGAAGG	180	
Db	121	CGCAAAATCGCGCTCGACCCAGCATATACCAAGTTCCCTGATGTAGTTGGAAGG	180	
QY	181	ACAGGCATGGAACATCTCACAATGAAGCTGGTGAAGTCAGTACTATTCGCAAGGAATG	240	
Db	181	ACAGGCATGGAACATCTCACAATGAAGCTGGTGAAGTCAGTACTATTCGCAAGGAATG	240	
QY	241	TTCAACCGGACAAATAGATTTACTTGAAGCGGCAATCAAGCCAGTTTATGTTTGTATGGC	300	
Db	241	TTCAACCGGACAAATAGATTTACTTGAAGCGGCAATCAAGCCAGTTTATGTTTGTATGGC	300	
QY	301	AAGCCTCTTGATATGAGAAACAAGAGCTTCTTAAAGATACTCAAAAAGAGATGATGCA	360	
Db	301	AAGCCTCTTGATATGAGAAACAAGAGCTTCTTAAAGATACTCAAAAAGAGATGATGCA	360	

Db 301 AAGCCTCTGATATGAGAAACAAGAGCTTGCTAAAGATCTCAAAAGAGATGATGCA 360
QY 361 ACCAAAGATCTGACTGAGGACGTAGAGTAGGAGATAAAGATGCGATTGAAAAATTGAGC 420
Db 361 ACCAAAGATCTGACTGAGGACGTAGAGTAGGAGATAAAGATGCGATTGAAAAATTGAGC 420
QY 421 AAGAGACTGTTAAAGTTCACAGGCACACACAGCAAGATGTTAAACGGCTATTAGACTT 480
Db 421 AAGAGACTGTTAAAGTTCACAGGCACACACAGCAAGATGTTAAACGGCTATTAGACTT 480
QY 481 ATGGGGTTCCTGTTCTAGAGGACCTCTGAAGCAGAGCAAGATGTCACGCCCTTTGC 540
Db 481 ATGGGGTTCCTGTTCTAGAGGACCTCTGAAGCAGAGCAAGATGTCACGCCCTTTGC 540
QY 541 ATAAACGATAAGGTGTCGCTGTTGCTCAGAGATATGGACTCCCTTTACTTTGGGGCT 600
Db 541 ATAAACGATAAGGTGTCGCTGTTGCTCAGAGATATGGACTCCCTTTACTTTGGGGCT 600
QY 601 CCAGGTTCCCTCGTCATTATGAGGATCAAGTTCACAGAAATACCTCTGATGGAATTT 660
Db 601 CCAGGTTCCCTCGTCATTATGAGGATCAAGTTCACAGAAATACCTCTGATGGAATTT 660
QY 661 GATGTTGCCAAGGTTTGGAGGACTTGAACCTCACCATGGACAGTTCATGATTTGTC 720
Db 661 GATGTTGCCAAGGTTTGGAGGACTTGAACCTCACCATGGACAGTTCATGATTTGTC 720
QY 721 ATCCTGTGTGATGTGACTATTGTGATAGCATCAAGGTTATCGGGGGGCAACAGCTCTG 780
Db 721 ATCCTGTGTGATGTGACTATTGTGATAGCATCAAGGTTATCGGGGGGCAACAGCTCTG 780
QY 781 AAATTTATTCGTCACATGGTCCATGAAAGCATCTTGGAGAACTCTTAATAAGACAGA 840
Db 781 AAATTTATTCGTCACATGGTCCATGAAAGCATCTTGGAGAACTCTTAATAAGACAGA 840
QY 841 TATCAAAATTCCTGAGGACTGGCCTTACCAAGAGCTCGAGCTTGTTCAGGAGCCTAAT 900
Db 841 TATCAAAATTCCTGAGGACTGGCCTTACCAAGAGCTCGAGCTTGTTCAGGAGCCTAAT 900
QY 901 GTCACATTTGATATTCCTGAGCTAAATGACCTGCACCTGATGAGGAGGCTCTCATAGT 960
Db 901 GTCACATTTGATATTCCTGAGCTAAATGACCTGCACCTGATGAGGAGGCTCTCATAGT 960
QY 961 TTCCTGTTAAAGATAATGTTTCAACGAAGATCGGTTGACAAAGGCCATAGAGAATC 1020
Db 961 TTCCTGTTAAAGATAATGTTTCAACGAAGATCGGTTGACAAAGGCCATAGAGAATC 1020
QY 1021 AAATCTGCCAAGATAAATCGTCCAGAGAGACTCGAGTCCCTTTTCAAGCCCACTGCC 1080
Db 1021 AAATCTGCCAAGATAAATCGTCCAGAGAGACTCGAGTCCCTTTTCAAGCCCACTGCC 1080
QY 1081 ACCACATCAGCAGCCTAAACAGGAAGAGACTTCGGATATAACAAAGCAGGAGCTGCC 1140
Db 1081 ACCACATCAGCAGCCTAAACAGGAAGAGACTTCGGATATAACAAAGCAGGAGCTGCC 1140
QY 1141 AACAAAGAAACAAAGGCTGTTGGAAGAAAGAAATAATCTTGGATGCTTGATGACAACTA 1200
Db 1141 AACAAAGAAACAAAGGCTGTTGGAAGAAAGAAATAATCTTGGATGCTTGATGACAACTA 1200
QY 1201 CGACTACAAAGACAGCGGTGCGGTGATCAGCTTCGCTTAGATTTAACTCCCTGTTTTA 1260
Db 1201 CGACTACAAAGACAGCGGTGCGGTGATCAGCTTCGCTTAGATTTAACTCCCTGTTTTA 1260
QY 1261 ACTCAGAGCTTTGGTAAAGTTCGCCATGTTTCAAGCTGGGGTAAAGTTAGTTGTGTTG 1320
Db 1261 ACTCAGAGCTTTGGTAAAGTTCGCCATGTTTCAAGCTGGGGTAAAGTTAGTTGTGTTG 1320
QY 1321 AAGAGATTGGTGATCAAGATTAACAAACTTATCGCTGTTTTTGAAGAAAAAATAAAAA 1380
Db 1321 AAGAGATTGGTGATCAAGATTAACAAACTTATCGCTGTTTTTGAAGAAAAAATAAAAA 1380
QY 1381 A 1381
Db 1381 A 1381

RESULT 2
US-09-426-557-1
; Sequence 1, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1463
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (85)...(1221)
US-09-426-557-1

Query Match 91.7%; Score 1267; DB 4; Length 1463;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1267; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 GCCACAGCCGCGCAGACGAGATGGGCATCAAGGTTTGGAGAACTGCTGGCGGACAAT 75
Db 64 GCCACAGCCGCGCAGACGAGATGGGCATCAAGGTTTGGAGAACTGCTGGCGGACAAT 123
QY 76 GCGCCCAAGGCGATGAAGGAGCAAGTTCGAGAGCTACTTCGGCCGCAAAATCGCCGTC 135
Db 124 GCGCCCAAGGCGATGAAGGAGCAAGTTCGAGAGCTACTTCGGCCGCAAAATCGCCGTC 183
QY 136 GAGCGCAGCATGAGCATATACCAGTTCCTGATTTGATGTTGGAAGCAGAGCATGGAAC 195
Db 184 GAGCGCAGCATGAGCATATACCAGTTCCTGATTTGATGTTGGAAGCAGAGCATGGAAC 243
QY 196 CTCACAAATGAGCTGGTGAAGTCACTAGTCACTTTCGAAGGAATGTTCAACCGGACAATA 255
Db 244 CTCACAAATGAGCTGGTGAAGTCACTAGTCACTTTCGAAGGAATGTTCAACCGGACAATA 303
QY 256 AGATTACTGGAAGCGGGAATCAAGCCAGTTCCTGATTTGATGTTGGAAGCAGCTCTGATG 315
Db 304 AGATTACTGGAAGCGGGAATCAAGCCAGTTCCTGATTTGATGTTGGAAGCAGCTCTGATG 363
QY 316 AAGAAACAAGAGCTTGCTAAAAGATCTCAAAAGAGATGATGCAACCAAGATCTGACT 375
Db 364 AAGAAACAAGAGCTTGCTAAAAGATCTCAAAAGAGATGATGCAACCAAGATCTGACT 423
QY 376 GAGGAGTAGAGTAGGAGATAAAGATGCGATTGAAAATTTGAGCAAGAGGACTGTAAG 435
Db 424 GAGGAGTAGAGTAGGAGATAAAGATGCGATTGAAAATTTGAGCAAGAGGACTGTAAG 483
QY 436 GTCACAAGGCAACACACACCAAGATTGTAACCGCTATTAAAGACTTATGGGGTTCCCTGTT 495
Db 484 GTCACAAGGCAACACACACCAAGATTGTAACCGCTATTAAAGACTTATGGGGTTCCCTGTT 543
QY 496 GTAGAGGACCTTCTGACAGCAGAGCAAGATGTCAGCCCTTTCGATTAACCATTAAGGTG 555
Db 544 GTAGAGGACCTTCTGACAGCAGAGCAAGATGTCAGCCCTTTCGATTAACCATTAAGGTG 603
QY 556 TTCGCTGTTGCTCAGAAGATATGGACTCCCTTACTTTTGGGGCTCCAGCGTTCCTTCGT 615
Db 604 TTCGCTGTTGCTCAGAAGATATGGACTCCCTTACTTTTGGGGCTCCAGCGTTCCTTCGT 663
QY 616 CATTTAATGATGATCAAGTTCCAGAAAAATACCTGTGATGGAATTTGATGTTCCCAAGTT 675
Db 664 CATTTAATGATGATCAAGTTCCAGAAAAATACCTGTGATGGAATTTGATGTTCCCAAGTT 723

```
QY 676 TTGGAGGAGCTTGAACTCACCATGACAGCTTCATTGATTGTGCATCCTGTGGATGT 735
DB 724 TTGGAGGAGCTTGAACTCACCATGACAGCTTCATTGATTGTGCATCCTGTGGATGT 783
QY 736 GACTATTGTGATAGCATCAAGGTATCGGGGGGCAACAGCTCTGAAACITATTTCGTCAA 795
DB 784 GACTATTGTGATAGCATCAAGGTATCGGGGGGCAACAGCTCTGAAACITATTTCGTCAA 843
QY 796 CATGGGTCCATAGAAGCATCTTGAGAAATCTTAATAAGACAGATATCAAAATTCCTGAG 855
DB 844 CATGGGTCCATAGAAGCATCTTGAGAAATCTTAATAAGACAGATATCAAAATTCCTGAG 903
QY 856 GACTGGCCTTACCAAGAAGCTCGAGCTTTGTTCAAGGAGCCTAATGTCACATTTGGATATT 915
DB 904 GACTGGCCTTACCAAGAAGCTCGAGCTTTGTTCAAGGAGCCTAATGTCACATTTGGATATT 963
QY 916 CCTGAGCTAAATGGACTGCACCTGATGAGGAGGTCTCTATAAGTTTCCTGGTAAAGAT 975
DB 964 CCTGAGCTAAATGGACTGCACCTGATGAGGAGGTCTCTATAAGTTTCCTGGTAAAGAT 1023
QY 976 AATGGTTCAACGAAGATCGGGTGACAAAGGCCATAGAGAAGATCAAAATTCGCCAAGAT 1035
DB 1024 AATGGTTTCAACGAAGATCGGGTGACAAAGGCCATAGAGAAGATCAAAATTCGCCAAGAT 1083
QY 1036 AATCGTCCCAAGGAAGACTCGAGTCCCTTTTCAAGCCACTGCCACACATCAGCACCG 1095
DB 1084 AATCGTCCCAAGGAAGACTCGAGTCCCTTTTCAAGCCACTGCCACACATCAGCACCG 1143
QY 1096 CTAAACGGAAGGAGACTTCGGATATAAACAAAGCAAGCAGCTGCGAACAAGAAACAAAG 1155
DB 1144 CTAAACGGAAGGAGACTTCGGATATAAACAAAGCAAGCAGCTGCGAACAAGAAACAAAG 1203
QY 1156 GCTGGTGAAGAAAGAAATACTTGGATGCTGTACACTACGACTACGAAAGCAG 1215
DB 1204 GCTGGTGAAGAAAGAAATACTTGGATGCTGTACACTACGACTACGAAAGCAG 1263
QY 1216 CGGTGGCGTGATCACTTCGCTTAGATTATTTAACTCCCTGTTTTAACTCAGAGCTTTGGT 1275
DB 1264 CGGTGGCGTGATCACTTCGCTTAGATTATTTAACTCCCTGTTTTAACTCAGAGCTTTGGT 1323
QY 1276 AAAAGTT 1282
DB 1324 AAAAGTT 1330
```

RESULT 3

```
US-09-426-557-3
; Sequence 3, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; FILE OF INVENTION: Thereof
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426.557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)...(1215)
US-09-426-557-3
```

```
Query Match      88.1%; Score 1216; DB 4; Length 1541;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1266; Conservative 0; Mismatches 1; Indels 0; Caps 0;
```

```
QY 16 GCCACAGCCGCCGAGAGATGCGCATCAAGGGTTTGGACAAACTGCTGCGGACAAT 75
DB 58 GCCACAGCCGCCGAGAGATGCGCATCAAGGGTTTGGACAAACTGCTGCGGACAAT 117
QY 76 GGGCCCAAGGCGATGAAGGAGCAGAAGTTCGAGAGCTACTTCGGCCGCCAAAATTCGGCGTC 135
DB 118 GGGCCCAAGGCGATGAAGGAGCAGAAGTTCGAGAGCTACTTCGGCCGCCAAAATTCGGCGTC 177
QY 136 GACGCGACATGAGCATATACCAGTTCCTGATTCAGTTCGAAAGGACAGCATGGAACAT 195
DB 178 GACGCGACATGAGCATATACCAGTTCCTGATTCAGTTCGAAAGGACAGCATGGAACAT 237
QY 196 CTCACAAATGAAGCTGGTGAAGTCACATGATCTTGAAGGAATGTTCAACCCGACAATA 255
DB 238 CTCACAAATGAAGCTGGTGAAGTCACATGATCTTGAAGGAATGTTCAACCCGACAATA 297
QY 256 AGATTACTGGAGGCGGGAATCAAGCCAGTTTATCTTTTGGATGCAAGCCTCCTGATATG 315
DB 298 AGATTACTGGAGGCGGGAATCAAGCCAGTTTATCTTTTGGATGCAAGCCTCCTGATATG 357
QY 316 AAGAAACAAGAGCTTGCTTAAAGATACCTCAAAAAGAGATGATGCAACCAAGATCTGACT 375
DB 358 AAGAAACAAGAGCTTGCTTAAAGATACCTCAAAAAGAGATGATGCAACCAAGATCTGACT 417
QY 376 GAGGCAGTAGAGGTAGGAGATAAGATCCGATTCGAAATTTGACAAAGAGACTGTAAAG 435
DB 418 GAGGCAGTAGAGGTAGGAGATAAGATCCGATTCGAAATTTGACAAAGAGACTGTAAAG 477
QY 436 GTCAACAAGGCAACCAAGAGATTTAAAGCGCTATTAAAGACTTATGGGGGTTCCTGTT 495
DB 478 GTCAACAAGGCAACCAAGAGATTTAAAGCGCTATTAAAGACTTATGGGGGTTCCTGTT 537
QY 496 GTAGAGGCACCTTCTGAGCAGAGAGCAGAAATGTCAGCCCTTTGCATTAACAGATAAGGTG 555
DB 538 GTAGAGGCACCTTCTGAGCAGAGAGCAGAAATGTCAGCCCTTTGCATTAACAGATAAGGTG 597
QY 556 TTTCGCTGTTGGTTTCAAGAAGATATGAGCTCCCTTACTTTTGGGGCTCCACCGTTCCCTTCGT 615
DB 598 TTTCGCTGTTGGTTTCAAGAAGATATGAGCTCCCTTACTTTTGGGGCTCCACCGTTCCCTTCGT 657
QY 616 CATTATTAATGGATCCCAAGTTTCAAGAAATACCTGTGATGGAATTTGATGTCGAAGTTT 675
DB 658 CATTATTAATGGATCCCAAGTTTCAAGAAATACCTGTGATGGAATTTGATGTCGAAGTTT 717
QY 676 TTGGAGGAGCTTGAACCTCACCATGAGCAGTTTCATTTGATTTGTGCATCCTGTGTGATGT 735
DB 718 TTGGAGGAGCTTGAACCTCACCATGAGCAGTTTCATTTGATTTGTGCATCCTGTGTGATGT 777
QY 736 GACTATTGTGATAGCATCAAGGTATCGGGGGGCAACAGCTCTGAAACITATTTCGTCAA 795
DB 778 GACTATTGTGATAGCATCAAGGTATCGGGGGGCAACAGCTCTGAAACITATTTCGTCAA 837
QY 796 CATGGTCCATAGAAGCATCTTGAGAAATCTTAATAAGACAGATATCAAAATTCCTGAG 855
DB 838 CATGGTCCATAGAAGCATCTTGAGAAATCTTAATAAGACAGATATCAAAATTCCTGAG 897
QY 856 GACTGGCCTTACCAAGAAGCTCGAGCCTTTGTTCAAGGAGCCTAATGTCACATTTGGATATT 915
DB 898 GACTGGCCTTACCAAGAAGCTCGAGCCTTTGTTCAAGGAGCCTAATGTCACATTTGGATATT 957
QY 916 CCTGAGCTAAATGGACTGCACCTGATGAGGAGGTCTCTATAAGTTTCCTGGTAAAGAT 975
DB 958 CCTGAGCTAAATGGACTGCACCTGATGAGGAGGTCTCTATAAGTTTCCTGGTAAAGAT 1017
QY 976 AATGGTTTCAACGAAGATCGGGTGACAAAGGCCATAGAGAAGATCAAAATTCGCCAAGAT 1035
DB 1018 AATGGTTTCAACGAAGATCGGGTGACAAAGGCCATAGAGAAGATCAAAATTCGCCAAGAT 1077
QY 1036 AATCGTCCCAAGGAAGACTCGAGTCCCTTTTCAAGCCACTGCCACACATCAGCACCG 1095
DB 1078 AATCGTCCCAAGGAAGACTCGAGTCCCTTTTCAAGCCACTGCCACACATCAGCACCG 1137
```

QY 1096 CTAAACGGAAGGAGACTTCGGATAAACAAAGCAGGAGCTGGCAACAAGAAACAAG 1155
Db 1138 CTAAACGGAAGGAGACTTCGGATAAACAAAGCAGGAGCTGGCAACAAGAAACAAG 1197
QY 1156 GCTGGTGAAGAAAGAAATAATCTTGGATGCTTGATGTACAACTACGACTACGAAAGCAG 1215
Db 1198 GCTGGTGAAGAAAGAAATAATCTTGGATGCTTGATGTACAACTACGACTACGAAAGCAG 1257
QY 1216 CGGTGCGTGATCACTTCGCTTAGATTAATTAACCTCCCTGTTTAACTCAGAGCTTTGGT 1275
Db 1258 CGGTGCGTGATCACTTCGCTTAGATTAATTAACCTCCCTGTTTAACTCAGAGCTTTGGT 1317
QY 1276 AAAAGTT 1282
Db 1318 AAAAGTT 1324

RESULT 4
US-09-426-557-7
; Sequence 7, Application US/09426557
; Patent No. 6232527
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Rad2/FEN-1 Orthologues and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0961
; CURRENT APPLICATION NUMBER: US/09/426,557
; CURRENT FILING DATE: 1999-10-22
; EARLIER APPLICATION NUMBER: 60/112,332
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1478
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (97)...(1233)
US-09-426-557-7

Query Match 61.5%; Score 850; DB 4; Length 1478;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1200; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 16 GCCACACGCCCGCAGACGAGATGGGCATCAAGGTTTGACGAACTGCTGGCGGACAAT 75
Db 76 GCCACACGCCCGCAGACGAGATGGGCATCAAGGTTTGACGAACTGCTGGCGGACAAT 135
QY 76 GCGCCCAAGCGGATGAAGGACGAGAGTTCCAGAGCTACTTCGCGCGCAAAATCGCCGTC 135
Db 136 GCGCCCAAGCGGATGAAGGACGAGAGTTCCAGAGCTACTTCGCGCGCAAAATCGCCGTC 195
QY 136 GACGCCAGCATGACATATACCACTTCCCTGATGTAGTTGGAAGGACAGGCATGGAAC 195
Db 196 GACGCCAGCATGACATATACCACTTCCCTGATGTAGTTGGAAGGACAGGCATGGAAC 255
QY 196 CTCACAATGAAGCTGGTGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACAATA 255
Db 256 CTCACAATGAAGCTGGTGAAGTCACTAGTCAATTTGCAAGGAATGTTCAACCGGACAATA 315
QY 256 AGATTACTGGAAGCGGGAATCAAGCCAGTTATGTTTGTGATGGAAGCCCTCTGTATG 315
Db 316 AGATTACTGGAAGCGGGAATCAAGCCAGTTATGTTTGTGATGGAAGCCCTCTGTATG 375
QY 316 AAGAACAAAGAGCTTGTAAAGATCTCAAAAGAGATGATGCAACCAAGAGTCTGACT 375
Db 376 AAGAACAAAGAGCTTGTAAAGATCTCAAAAGAGATGATGCAACCAAGAGTCTGACT 435
QY 376 GAGGAGTAGAGGTAGGAGATAAGATGCGATTGAAATTTGACAAAGAGAGCTGTAAG 435
Db 436 GAGGAGTAGAGGTAGGAGATAAGATGCGATTGAAATTTGACAAAGAGAGCTGTAAG 495

QY 436 GTCACAGGCAACACAAAGGATTTAAACGGCTATTAAAGACTTATGGGGTTCCTGTT 495
Db 496 GTCACAGGCAACACAAAGGATTTAAACGGCTATTAAAGACTTATGGGGTTCCTGTT 555
QY 496 GTAGAGCACCTTCTGAAGCAGAAAGCAGATGTGCAGCCCTTTGCAATAACGATAAGGTG 555
Db 556 GTAGAGCACCTTCTGAAGCAGAAAGCAGATGTGCAGCCCTTTGCAATAACGATAAGGTG 615
QY 556 TTCCTGCTTCTCAGAGATATGGACTCCCTTACTTTTGGGGCTCCACGGTTCCCTGTT 615
Db 616 TTCCTGCTTCTCAGAGATATGGACTCCCTTACTTTTGGGGCTCCACGGTTCCCTGTT 675
QY 616 CATTTAATGATCCAACTTCCAAAGAAATACCTGTGATGGAATTTGATGTTGCCAAGGTT 675
Db 676 CATTTAATGATCCAACTTCCAAAGAAATACCTGTGATGGAATTTGATGTTGCCAAGGTT 735
QY 676 TTGAGGAGCTTGAACCTCACCATGGACCACTTCATTGATTTTGGATCTCTGTGATGTT 735
Db 736 TTGAGGAGCTTGAACCTCACCATGGACCACTTCATTGATTTTGGATCTCTGTGATGTT 795
QY 736 GACTATTGTGATAGCATCAAGGTTATCGGGGGCAACAGCTCTGAAACTTATTTCGTCAA 795
Db 796 GACTATTGTGATAGCATCAAGGTTATCGGGGGCAACAGCTCTGAAACTTATTTCGTCAA 855
QY 796 CATGGTCCATAGAAAGCATCTTGGAGAAATCTTAATAAGACAGATATCAAAATTCCTGAG 855
Db 856 CATGGTCCATAGAAAGCATCTTGGAGAAATCTTAATAAGACAGATATCAAAATTCCTGAG 915
QY 856 GACTGGCTTTACCAAGAGCTCGAGCTTGTTCAGAGGAGCTAATGTCACATTTGGATATT 915
Db 916 GACTGGCTTTACCAAGAGCTCGAGCTTGTTCAGAGGAGCTAATGTCACATTTGGATATT 975
QY 916 CCTGAGCTAAAATGACTGCACTGTGAGGAGGCTCTCAATGTTTCTGGTAAAAGAT 975
Db 976 CCTGAGCTAAAATGACTGCACTGTGAGGAGGCTCTCAATGTTTCTGGTAAAAGAT 1035
QY 976 AATGTTTTCAACGAAGATCGGGTGCACAAAGCCATAGAGAAGTCAATCTGCCAAGAAT 1035
Db 1036 AATGTTTTCAACGAAGATCGGGTGCACAAAGCCATAGAGAAGTCAATCTGCCAAGAAT 1095
QY 1036 AATCGTCGCAAGGAGACTCGAGTCTTTTCAAGGCAACTGCCACCATCAGCACCG 1095
Db 1096 AATCGTCGCAAGGAGACTCGAGTCTTTTCAAGGCAACTGCCACCATCAGCACCG 1155
QY 1096 CTAAACGGAAGGAGACTTCGGATAAACAAAGCAGGAGCTGGCAACAAGAAACAAG 1155
Db 1156 CTAAACGGAAGGAGACTTCGGATAAACAAAGCAGGAGCTGGCAACAAGAAACAAG 1215
QY 1156 GCTGTGGAAGAAAGAAATAATCTTGGATGCTTGATGTACAACTACGACTACGAAAGCAG 1215
Db 1216 GCTGTGGAAGAAAGAAATAATCTTGGATGCTTGATGTACAACTACGACTACGAAAGCAG 1275
QY 1216 CGGTGGC 1222
Db 1276 CGGTGGC 1282

RESULT 5
US-09-136-073-1
; Sequence 1, Application US/09136073
; Patent No. 6043076
; GENERAL INFORMATION:
; APPLICANT: Ramakrishnan, Santha
; TITLE OF INVENTION: GENE ENCODING 2,3-DIHYDROXYBENZOIC ACID DECARBOXYLASE
; FILE REFERENCE: UNV.53687
; CURRENT APPLICATION NUMBER: US/09/136,073
; CURRENT FILING DATE: 1998-08-18
; EARLIER APPLICATION NUMBER: 60/056,621
; EARLIER FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Corel Wordperfect 6.1 for Windows
; SEQ ID NO 1
; LENGTH: 1096

```

; TYPE: DNA
; ORGANISM: Aspergillus niger
; FEATURE:
; NAME/KEY: gene
; LOCATION: 1-1096
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1-1029
; FEATURE:
; NAME/KEY: polyA_site
; LOCATION: 1073
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: 1056-1061
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: 1027-1073
; FEATURE:
; NAME/KEY: source/Aspergillus niger
; LOCATION: 1-1096
; PUBLICATION INFORMATION:
; AUTHORS: Santha, Ramakrishnan
; AUTHORS: Dickman, Martin B.
; AUTHORS: O'Leary, Marion H.
; TITLE: 2,3-Dihydroxybenzoic Acid Decarboxylase From Aspergillus niger:
; TITLE: Mechanism, Cloning And Overexpression.
; JOURNAL: Faseb Journal
; VOLUME: 11
; ISSUE: 9
; PAGES: A1017
; DATE: 1997-07-31
US-09-136-073-1

```

```

Query Match      2.0%; Score 27; DB 3; Length 1096;
Best Local Similarity 100.0%; Pred. No. 0.014; Indels 0; Gaps 0;
Matches 27; Conservative 0; Mismatches 0;

```

```

QY 1355 CTGTTTTTTGAAAAAAGAAAAA 1381
      |||||||
DB 1067 CTGTTTTTTGAAAAAAGAAAAA 1093

```

```

RESULT 6
US-08-808-931-9
; Sequence 9, Application US/08808931
; Patent No. 5939602
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; TITLE OF INVENTION: DNA Molecules Encoding Plant
; TITLE OF INVENTION: Protoporphyrogen Oxidase and Inhibitor-Resistant Mutants
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5939602artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,931
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1811 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum (wheat)
; IMMEDIATE SOURCE:
; CLONE: pWDC-13 (NRRL B-21545)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1589
; OTHER INFORMATION: /product= "wheat protox-1"
US-08-808-931-9

```

```

Query Match      1.7%; Score 24; DB 2; Length 1811;
Best Local Similarity 100.0%; Pred. No. 0.25; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;

```

```

QY 1358 TTTTGTGAAAAAAGAAAAA 1381
      |||||||
DB 1786 TTTTGTGAAAAAAGAAAAA 1809

```

```

RESULT 7
US-08-808-923-9
; Sequence 9, Application US/08808323
; Patent No. 6018105
; GENERAL INFORMATION:
; APPLICANT: Johnson, Marie
; APPLICANT: Volrath, Sandra
; APPLICANT: Ward, Eric
; TITLE OF INVENTION: Promoters from Plant
; TITLE OF INVENTION: Protoporphyrogen Oxidase Genes
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6018105artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,323
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1846
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1811 base pairs
; TYPE: nucleic acid
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORGANISM: Triticum aestivum (wheat)
; IMMEDIATE SOURCE:
; CLONE: pWDC-13 (NRRL B-21545)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1589
; OTHER INFORMATION: /product= "wheat protox-1"
;
US-08-808-323-9

Query Match 1.7%; Score 24; DB 3; Length 1811;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1358 TTTTGTGAAAAAAGAAAAA 1381
Db 1786 TTTTGTGAAAAAAGAAAAA 1809

RESULT 8
US-09-050-603A-9
; Sequence 9, Application US/09050603A
; Patent No. 6023012
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Heifetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding plant
; TITLE OF INVENTION: Protoporphyrinogen Oxidase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6023012artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,603A
; FILING DATE: 30-MAR-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/808,931
; FILING DATE: 28-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705

```

```

; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1811 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum (wheat)
; IMMEDIATE SOURCE:
; CLONE: pWDC-13 (NRRL B-21545)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1589
; OTHER INFORMATION: /product= "wheat protox-1"
;
US-09-050-603A-9

Query Match 1.7%; Score 24; DB 3; Length 1811;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1358 TTTTGTGAAAAAAGAAAAA 1381
Db 1786 TTTTGTGAAAAAAGAAAAA 1809

RESULT 9
US-09-102-420B-9
; Sequence 9, Application US/09102420B
; Patent No. 6084155
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Ward, Eric
; APPLICANT: Heifetz, Peter
; TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
; TITLE OF INVENTION: OXIDASE ("PROTOX")
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6084155artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/102,420B
; FILING DATE: 22-JUN-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/059,164
; FILING DATE: 13-APR-1998
; PRIOR APPLICATION DATA:

```



```
; APPLICATION NUMBER: US 09/050,603
; FILING DATE: 30-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/126,430
; FILING DATE: 11-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/808,931
; FILING DATE: 28-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/472,028
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1811 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum (wheat)
; IMMEDIATE SOURCE:
; CLONE: pWDC-13 (NRRL B-21545)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1589
; OTHER INFORMATION: /product= "wheat protox-1"
;
US-09-102-4208-9

Query Match 1.7%; Score 24; DB 3; Length 1811;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAA 1381
Db 1786 TTTTGTGAAAAA 1809

RESULT 10
US-09-497-698-9
; Sequence 9, Application US/09497698
; Patent No. 6308458
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; Johnson, Marie
; Ward, Eric
; Heifetz, Peter
; TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
; OXIDASE ("PROTOX")
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6308458artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709

; APPLICATION NUMBER: US 09/050,603
; FILING DATE: 30-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/126,430
; FILING DATE: 11-MAR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/808,931
; FILING DATE: 28-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/472,028
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1811 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Triticum aestivum (wheat)
; IMMEDIATE SOURCE:
; CLONE: pWDC-13 (NRRL B-21545)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1589
; OTHER INFORMATION: /product= "wheat protox-1"
;
US-09-102-4208-9

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/497,698
FILING DATE: 03-Feb-2000
CLASSIFICATION: <Unknown>
30-MAR-1998
11-MAR-1998
28-FEB-1997
28-FEB-1996
28-FEB-1996
21-JUN-1996
06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/102,420
FILING DATE: <Unknown>
APPLICATION NUMBER: US 09/050,603
FILING DATE: 30-MAR-1998
APPLICATION NUMBER: US 60/126,430
FILING DATE: 11-MAR-1998
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
APPLICATION NUMBER: US 08/472,028
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1811 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORGANISM: Triticum aestivum (wheat)
IMMEDIATE SOURCE:
CLONE: pWDC-13 (NRRL B-21545)
FEATURE:
NAME/KEY: CDS
LOCATION: 3..1589
OTHER INFORMATION: /product= "wheat protox-1"
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-497-698-9

Query Match 1.7%; Score 24; DB 4; Length 1811;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1358 TTTTGTGAAAAA 1381
Db 1786 TTTTGTGAAAAA 1809

RESULT 11
US-09-234-332-5/C
; Sequence 5, Application US/09234332A
; Patent No. 6087168
; GENERAL INFORMATION:
```

APPLICANT: Cedars-Sinai Medical Center
APPLICANT: Michel F. Levesque, M.D.
APPLICANT: Toomas Neuman, Ph.D.
TITLE OF INVENTION: CONVERSION OF NON-NEURONAL CELLS INTO
FILE REFERENCE: P07 41494
CURRENT APPLICATION NUMBER: US/09/234,332A
CURRENT FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 3138
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: gene
LOCATION: (0)...(0)
OTHER INFORMATION: Zic 1 Protein gene; Genbank Accession D76435
US-09-234-332-5

Query Match 1.7%; Score 23; DB 3; Length 3138;
Best Local Similarity 100.0%; Pred. No. 0.62; Mismatches 0; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1359 TTTTGAAGAAAAA 1381
DB 218 TTTTGAAGAAAAA 196

RESULT 12
US-08-865-273-1
Sequence 1, Application US/08865273
Patent No. 5954100
GENERAL INFORMATION:
APPLICANT: ZHU, YUAN
APPLICANT: NAMBI, PONNAL
APPLICANT: PULLEN, MARK A
TITLE OF INVENTION: NOVEL HAS2 SPLICING VARIANT
INFLAMMATORY DISEASES AND MYOCARDIAL ISCHEMIA
HOEFC11: A TARGET IN CHRONIC RENAL FAILURE,
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: RAYNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/865,273
FILING DATE: 29-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70053
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1051 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: cDNA
US-08-865-273-1

Query Match 1.6%; Score 22; DB 2; Length 1051;
Best Local Similarity 100.0%; Pred. No. 1.8; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1360 TTTTGAAGAAAAA 1381
DB 1030 TTTTGAAGAAAAA 1051

RESULT 13
US-09-385-174-1
Sequence 1, Application US/09385174
Patent No. 6350446
GENERAL INFORMATION:
APPLICANT: ZHU, YUAN
APPLICANT: NAMBI, PONNAL
APPLICANT: PULLEN, MARK A
TITLE OF INVENTION: NOVEL HAS2 SPLICING VARIANT
INFLAMMATORY DISEASES AND MYOCARDIAL ISCHEMIA
HOEFC11: A TARGET IN CHRONIC RENAL FAILURE,
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: RAYNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/385,174
FILING DATE: 30-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/865,273
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70053
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1051 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-385-174-1

Query Match 1.6%; Score 22; DB 4; Length 1051;
Best Local Similarity 100.0%; Pred. No. 1.8; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1360 TTTTGAAGAAAAA 1381
DB 1030 TTTTGAAGAAAAA 1051

RESULT 14
US-09-045-193-1
Sequence 1, Application US/09045193

```
; Patent No. 6245550
; GENERAL INFORMATION:
; APPLICANT: HENSLEY, PRESTON
; APPLICANT: ROSE, GEORGE
; APPLICANT: AURORA, RAJEV
; APPLICANT: ABDEL-MEGUID, SHERIN
; APPLICANT: YOUNG, PETER
; APPLICANT: MOONEY, JEFFREY
; APPLICANT: BERGSMAN, DICK
; APPLICANT: GUERRERA, STEPHANIE
; APPLICANT: ELLIS, CATHERINE
; TITLE OF INVENTION: The Cytokine Family Member
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,193
; FILING DATE: 20-MAR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GP-70421
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1067 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-045-193-1
;
; Query Match 1.6%; Score 22; DB 4; Length 1067;
; Best Local Similarity 100.0%; Pred. No. 1.8;
; Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1360 TTTTGAAGAAAAA 1381
; Db 1046 TTTTGAAGAAAAA 1067
;
; RESULT 15
; US-08-698-805-5
; Sequence 5, Application US/08698805
; Patent No. 5869288
; GENERAL INFORMATION:
; APPLICANT: Chapman, Martin
; APPLICANT: Arruda, L. Karla
; TITLE OF INVENTION: Molecular Cloning of Cockroach
; TITLE OF INVENTION: Allergens, Amino Acid and Nucleotide Sequences Therefore,
; TITLE OF INVENTION: and Recombinant Expression Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Obion, Spivak, McClelland, Maier & Neustadt,
```

```
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/698,805
; FILING DATE: 16-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/002,510
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelber, Steven B.
; REGISTRATION NUMBER: 30,073
; REFERENCE/DOCKET NUMBER: 494-203-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..602
; US-08-698-805-5
;
; Query Match 1.6%; Score 22; DB 2; Length 1140;
; Best Local Similarity 100.0%; Pred. No. 1.8;
; Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1360 TTTTGAAGAAAAA 1381
; Db 1117 TTTTGAAGAAAAA 1138
;
; RESULT 16
; US-09-347-798-1
; Sequence 1, Application US/09347798
; Patent No. 6242256
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Edgar B.
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Hitz, William D.
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Ornithine Biosynthesis Enzymes
; FILE REFERENCE: BB-1174-B
; CURRENT APPLICATION NUMBER: US/09/347,798
; CURRENT FILING DATE: 1999-07-02
; EARLIER APPLICATION NUMBER: 60/093,209
; EARLIER FILING DATE: July 17, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1193
; TYPE: DNA
; ORGANISM: Helianthus tuberosus
; US-09-347-798-1
;
; Query Match 1.6%; Score 22; DB 4; Length 1193;
; Best Local Similarity 100.0%; Pred. No. 1.8;
; Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1360 TTTTGAIAAAAAAAAAAAAA 1381
|||||
Db 1169 TTTTGAIAAAAAAAAAAAAA 1190

RESULT 17
US-08-820-170A-9
; Sequence 9, Application US/08820170A
; Patent No. 5831058
; GENERAL INFORMATION:
; APPLICANT: Tsutomu, FUJIWARA
; APPLICANT: Takeshi, WATANABE
; APPLICANT: Masato, HORIE
; APPLICANT: Toyomasa, KATAGIRI
; TITLE OF INVENTION: HUMAN GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/820,170A
; FILING DATE:
; CLASSIFICATION: 536
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1493 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA(genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; LIBRARY: Human fetal brain cDNA library
; CLONE: GEN-025F07
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 95..478

US-08-820-170A-9
Query Match 1.6%; Score 22; DB 2; Length 1493;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1360 TTTTGAIAAAAAAAAAAAAA 1381
|||||
Db 1471 TTTTGAIAAAAAAAAAAAAA 1492

RESULT 18
US-09-055-699-9
; Sequence 9, Application US/09055699
; Patent No. 6005088
; GENERAL INFORMATION:
; APPLICANT: Tsutomu, FUJIWARA
; APPLICANT: Takeshi, WATANABE
; APPLICANT: Masato, HORIE
; APPLICANT: Toyomasa, KATAGIRI
; TITLE OF INVENTION: HUMAN GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/820,170A
; FILING DATE:
; CLASSIFICATION: 536
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1493 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA(genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; LIBRARY: Human fetal brain cDNA library
; CLONE: GEN-025F07
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 95..478

US-08-820-170A-9
Query Match 1.6%; Score 22; DB 2; Length 1493;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1360 TTTTGAIAAAAAAAAAAAAA 1381
|||||
Db 1471 TTTTGAIAAAAAAAAAAAAA 1492

RESULT 19
US-09-273-565-9
; Sequence 9, Application US/09273565A
; Patent No. 6166190
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, TSUTOMU
; APPLICANT: WATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/273,565A
; CURRENT FILING DATE: 1999-03-22
; EARLIER APPLICATION NUMBER: 09/055,699
; EARLIER FILING DATE: 1998-04-07
; EARLIER APPLICATION NUMBER: 08/820,170
; EARLIER FILING DATE: 1997-03-19
; EARLIER APPLICATION NUMBER: JP 63410/1996
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: JP 69163/1997
; EARLIER FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/055,699
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/820,170
; FILING DATE:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1493 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA(genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; IMMEDIATE SOURCE:
; LIBRARY: Human fetal brain cDNA library
; CLONE: GEN-025F07
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 95..478
US-09-055-699-9

Query Match 1.6%; Score 22; DB 3; Length 1493;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1360 TTTTGAIAAAAAAAAAAAAA 1381
|||||
Db 1471 TTTTGAIAAAAAAAAAAAAA 1492

RESULT 19
US-09-273-565-9
; Sequence 9, Application US/09273565A
; Patent No. 6166190
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, TSUTOMU
; APPLICANT: WATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/273,565A
; CURRENT FILING DATE: 1999-03-22
; EARLIER APPLICATION NUMBER: 09/055,699
; EARLIER FILING DATE: 1998-04-07
; EARLIER APPLICATION NUMBER: 08/820,170
; EARLIER FILING DATE: 1997-03-19
; EARLIER APPLICATION NUMBER: JP 63410/1996
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: JP 69163/1997
; EARLIER FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9

```
; LENGTH: 1493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(478)
US-09-273-565-9

Query Match      1.6%  Score 22; DB 4; Length 1493;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1360 TTTTGAAGAAAAAAGAAAAA 1381
      |||||
Db 1471 TTTTGAAGAAAAAAGAAAAA 1492

RESULT 20
US-09-565-538-9
; Sequence 9, Application US/09565538
; Patent No. 6333404
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, TSUTOMU
; APPLICANT: WATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
; FILE REFERENCE: 0-53599
; CURRENT APPLICATION NUMBER: US/09/565,538
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 09/273,565
; PRIOR FILING DATE: 1999-03-22
; PRIOR APPLICATION NUMBER: 09/055,699
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: 08/820,170
; PRIOR FILING DATE: 1997-03-19
; PRIOR APPLICATION NUMBER: JP 63410/1996
; PRIOR FILING DATE: 1996-03-19
; PRIOR APPLICATION NUMBER: JP 69163/1997
; PRIOR FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(478)
US-09-565-538-9

Query Match      1.6%  Score 22; DB 4; Length 1493;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1360 TTTTGAAGAAAAAAGAAAAA 1381
      |||||
Db 1471 TTTTGAAGAAAAAAGAAAAA 1492

RESULT 21
US-08-464-523B-1
; Sequence 1, Application US/08464523B
; Patent No. 5723761
; GENERAL INFORMATION:
; APPLICANT: Toni A. Voelker
; APPLICANT: Ling Yuan
; APPLICANT: Jean Kridl
; APPLICANT: Deborah Hawkins
; APPLICANT: Aubrey Jones
; TITLE OF INVENTION: Plant Acyl ACP Thioesterase
; TITLE OF INVENTION: Sequences
; NUMBER OF SEQUENCES: 33
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1(a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,523B
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/13131
; FILING DATE: 10-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/152,004
; FILING DATE: 10-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/261,695
; FILING DATE: 16-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 100-1WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1745 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
US-08-464-523B-1

Query Match      1.6%  Score 22; DB 1; Length 1745;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1360 TTTTGAAGAAAAAAGAAAAA 1381
      |||||
Db 1720 TTTTGAAGAAAAAAGAAAAA 1741

RESULT 22
US-08-974-549A-4
; Sequence 4, Application US/08974549A
; Patent No. 6166178
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin B.
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit
; NUMBER OF SEQUENCES: 727
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/974,549A
  FILING DATE: 19-NOV-1997
  CLASSIFICATION: 536
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/724,643
    FILING DATE: 01-OCT-1996
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/844,419
    FILING DATE: 18-APR-1997
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/846,017
    FILING DATE: 25-APR-1997
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/851,843
    FILING DATE: 06-MAY-1997
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/854,050
    FILING DATE: 09-MAY-1997
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/911,312
    FILING DATE: 14-AUG-1997
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/912,951
    FILING DATE: 14-AUG-1997
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/915,503
    FILING DATE: 14-AUG-1997
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: WO PCT/US97/17618
    FILING DATE: 01-OCT-1997
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: WO PCT/US97/17885
    FILING DATE: 01-OCT-1997
  ATTORNEY/AGENT INFORMATION:
    NAME: Apple, Randolph Ted
    REGISTRATION NUMBER: 36,429
    REFERENCE/DOCKET NUMBER: 015389-0026100S
    TELEPHONE: (415) 576-0200
    TELEFAX: (415) 576-0300
    INFORMATION FOR SEQ ID NO: 4:
      SEQUENCE CHARACTERISTICS:
        LENGTH: 3855 base pairs
        TYPE: nucleic acid
        STRANDEDNESS: single
        TOPOLOGY: linear
        MOLECULE TYPE: cDNA
        FEATURE:
          NAME/KEY: -
          LOCATION: 1..3855
          OTHER INFORMATION: /note= "nucleic acid sequence with an
            OTHER INFORMATION: open reading frame encoding a delta-182
            OTHER INFORMATION: variant polypeptide"
          FEATURE:
            NAME/KEY: CDS
            LOCATION: 56..2479
            OTHER INFORMATION: /product= "delta-182 variant
            OTHER INFORMATION: polypeptide"
US-08-974-549A-4

```

```

Query Match 1.6%; Score 22; DB 4; Length 3855;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1360 TTTTGAAAAA...AAAAAAAAA 1381
Db 3822 TTTTGAAAAA...AAAAAAAAA 3843

```

```

RESULT 23
US-08-851-843A-173
: Sequence 173, Application US/08851843A
: Patent No. 6093809
: GENERAL INFORMATION:
: APPLICANT: Cech, Thomas R.
: APPLICANT: Lingner, Joachim
: APPLICANT: Nakamura, Toru
: APPLICANT: Chapman, Karen B.
: APPLICANT: Morin, Gregg B.
: APPLICANT: Harley, Calvin
: APPLICANT: Andrews, William H.
: TITLE OF INVENTION: No. 6093809el Telomerase
: NUMBER OF SEQUENCES: 225
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, 8th Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: United States of America
: ZIP: 94111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/851,843A
: FILING DATE: 06-MAY-1997
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/846,017
: FILING DATE: 25-APR-1997
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/844,419
: FILING DATE: 18-APR-1997
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/724,643
: FILING DATE: 01-OCT-1996
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Apple, Randolph T.
: REGISTRATION NUMBER: 36,429
: REFERENCE/DOCKET NUMBER: 015389-0029300S
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
: INFORMATION FOR SEQ ID NO: 173:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4029 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY:
: LOCATION: 1..4029
: OTHER INFORMATION: /note= "preliminary sequence for
: OTHER INFORMATION: human rrf cDNA insert of
: OTHER INFORMATION: plasmid pGRN121"
US-08-851-843A-173

```

```

Query Match 1.6%; Score 22; DB 3; Length 4029;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1360 TTTTGAAAAA...AAAAAAAAA 1381
Db 3999 TTTTGAAAAA...AAAAAAAAA 4020

```